A Predictive Model of Colour Differentiation

A Thesis Submitted to the
College of Graduate Studies and Research
in Partial Fulfillment of the Requirements
for the degree of Master of Science
in the Department of Computer Science
University of Saskatchewan
Saskatoon

By

David Raymond Flatla

Permission to Use

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Computer Science
176 Thorvaldson Building
110 Science Place
University of Saskatchewan
Saskatoon, Saskatchewan
Canada
S7N 5C9

Abstract

The ability to differentiate between colours varies from individual to individual. This variation is attributed to factors such as the presence of colour blindness. Colour is used to encode information in information visualizations. An example of such an encoding is categorization using colour (e.g., green for land, blue for water).

As a result of the variation in colour differentiation ability among individuals, many people experience difficulties when using colour-encoded information visualizations. These difficulties result from the inability to adequately differentiate between two colours, resulting in confusion, errors, frustration, and dissatisfaction.

If a user-specific model of colour differentiation was available, these difficulties could be predicted and corrected. Prediction and correction of these difficulties would reduce the amount of confusion, errors, frustration, and dissatisfaction experienced by users. This thesis presents a model of colour differentiation that is tuned to the abilities of a particular user. To construct this model, a series of judgement tasks are performed by the user. The data from these judgement tasks is used to calibrate a general colour differentiation model to the user. This calibrated model is used to construct a predictor. This predictor can then be used to make predictions about the user's ability to differentiate between two colours.

Two participant-based studies were used to evaluate this solution. The first study evaluated the basic approach used to model colour differentiation. The second study evaluated the accuracy of the predictor by comparing its performance to the performance of human participants. It was found that the predictor was as accurate as the human participants 86.3% of the time. Using such a predictor, the colour differentiation abilities of particular users can be accurately modeled.

Acknowledgements

I would like to thank Dr. Carl Gutwin for all of his helpful advice and his supervision of this project. I would also like to thank my wife, Amanda for her dedication and her willingness to help. I would also like to thank my committee, Dr. Mark Eramian, Dr. Gordon McCalla (chair), and Dr. W. J. (Chris) Zhang (external).

Dedicated to my wife, who helped me from beginning to end; my son, who always smiled when he saw me; and my parents, who have always supported me.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Information visualization is the visual representation of non-visual information. For example, a three-dimensional scatterplot of credit card transactions detailing the location, amount and time of each transaction is an information visualization. Although these visualizations can take many forms, they often employ colour as a means to visualize one or more dimensions or characteristics of the data being represented. Some examples of commonly used colour techniques are categorical encoding, popout, highlighting, and continuums. As an example of colour use, the NameVoyager system (Figure 1.1) visualizes several aspects of baby name popularity, including which names are for boys (blue) and which are for girls (pink). This also serves as an example of categorical encoding.

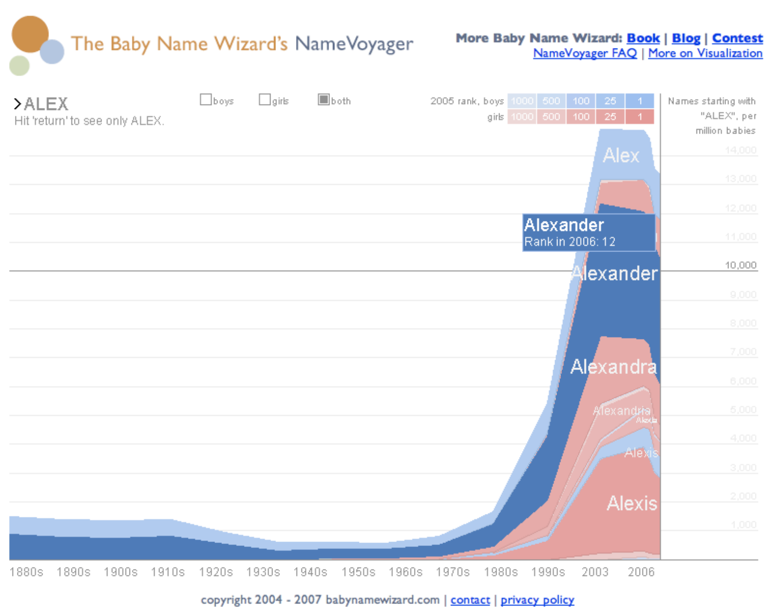

Figure 1.1: Example of colour use in information visualization. From: `http://www.babynamewizard.com/namevoyager/lnv0105.html`

Current information visualization techniques assume some degree of consistency between the colours seen by the designer of a visualization and what the observer sees. This allows the designer to assume that two colours which are differentiable for them will be differentiable for their audience. This is not the case with users of the visualization who are colour blind.

Colour blindness is an often-misunderstood term used to describe the visual perception of people with atypical colour perception. A colour blind individual is usually able to perceive colours, although in a limited fashion, and can be thought of as seeing a reduced set of colours compared to an individual with typical colour vision. Although the most common forms of colour blindness have genetic causes, there are also acquired forms of colour blindness resulting from prescription drug use, exposure to organic compounds, yellowing of the optical lens with age, and long-term diabetes. Colour blindness comes in many varieties and severities - ranging from individuals who never notice their colour blindness to people who have no colour perception beyond shades of grey.

Because a colour blind individual perceives a reduced set of colours, they tend to confuse colours which the remainder of the population perceives as different. This carries over to colour-utilizing information visualizations, and as a result, individuals with colour blindness often have difficulty using information visualizations.

There have been several proposed solutions to these difficulties. Each of these solutions adapts the colours used in a visualization to accommodate colour blind users. As an example, suppose a visualization uses red and green to encode two categories of information. If the colour blind individual has difficulty differentiating between red and green (i.e., they get them confused), the solution will recolour the image such that a new colour scheme is used (perhaps yellow and blue). In order to detect when a visualization uses colours that a colour blind individual will potentially confuse, each solution relies upon a model of colour perception. This model is also used to identify replacement colours that will not be confused by the user. These models, and the prediction of how well a user can differentiate colours, will be explored in this thesis.

As an example of how this model is used, suppose a visualization which uses RGB to denote colour is being analyzed for confusing colours. The individual for which this analysis is being performed happens to be missing all of the long-wavelength sensitive cones. The general process is:

1. Visualization colours are transformed using an orthogonal transformation into a colour encoding that represents colours as the amount of stimulation for the three types of cones in the typical human eye.

2. The new encoding is modified according to the particular form of colour blindness being modeled. In this case, this would involve setting the values for the long-wavelength cone to 0.

3. The modified encoding is then transformed back to RGB using the inverse of the orthogonal transformation used in Step 1.

4. The resulting visualization is analyzed to see if any colours that were different in the original visualization are now the same. These are the confused colours.

The accuracy of this model relies on the orthogonal transformation and its inverse, as well as the accuracy of the cone stimulation representation. To ensure the accuracy of this transformation and representation, the following conditions must to be met:

- details are available about the specific type and severity of colour blindness being accommodated

- all hardware and software are properly calibrated (operation system, video card, monitor, gamma corrections)

- sources of external lighting are held constant (generally at a bright white)

- the colour blindness being accommodated is only caused by retinal photosensor abnormalities, and not by higher-level cognitive impairment

- no population variation exists in the peak sensitivities of cones

These conditions, however, are often not achievable. Everyday computer use rarely involves correctly-calibrated graphics hardware and software colour calibration. Likewise, the everyday computer user does not use the computer in controlled lighting situations, where the amount and quality of incident light is regulated. Colour vision tests rarely identify the type and severity of colour blindness, but only detect the presence or absence of the disorder [5, 13]. Colour blindness can also be present in individuals with typical cones. In these cases, colour blindness may be caused by prescription drug use, exposure to toxic chemicals, disease, mental illness, aneurysm, or stroke, as these affect higher-level colour processing centers in the brain. [6, 10]. There is also significant variation in spectral response curves from person to person [18], meaning that the orthogonal transformation converts to an *idealized* cone stimulation representation, which is not necessarily representative of the particular individual under consideration.

In general, the problem with current colour-adaptation solutions is that their underlying models do not take into account a wide variety of these contextual factors.

1.1 Problem

The problem to be addressed in this thesis is: *Current colour perception models fail to accurately predict the colour confusion problems for a particular individual because they are not sensitive to the specific context of a colour perception task.*

The context of a colour perception task is the set of all factors which influence colour perception. These include the colour generating hardware, environmental lighting and reflection, the presence of colour blindness in the user, the presence of any filtering mediums which alter colour perception (such as tinted contact lenses), etc. In order for the current modeling technique described above to determine which colours are confusing and which are not, every potential colour-perception-modifying factor must either be known or controlled. This is the basis for the conditions listed above.

In a general computing environment, every factor that modifies the colour perception experience is rarely known or controlled. As the properties of the orthogonal transformation used in the modeling process described above are based upon these factors being controlled or known, the general computing environment invalidates this orthogonal transformation. As an example, let the orthogonal transformation be defined assuming the white light distribution of the monitor approximates D65[1] light [47]. If the white light distribution of the monitor being used does not approximate D65 light, then the orthogonal transformation will not properly convert the original colour representation into a cone stimulation representation. The resulting cone stimulation representation will have the stimulation incorrectly distributed over the three cone types because of the incorrect light distribution. When the cone representation is modified and converted back to the original colour representation, the colours will be different from the original colours, but because the transformation was not correct, some colours will be mistakenly identified as confused, even though they are not, and come colours will not be identified as confused, even though they are.

As an extension to this example, assume the individual for which the process is performed is colour blind, but has been incorrectly diagnosed as having a malformed medium-wavelength cone, when in actuality they have no medium-wavelength cones at all. In this case, the modification of the cone stimulation representation will retain some stimulation of the medium wavelength cones. When this is converted back to the original colour representation using the inverse orthogonal transformation, colours which the individual actually confuses will be not identified as such.

1.2 Motivation

Colour is readily available for information visualization designers. Computer monitors, televisions, MP3 players, cellular phones, and portable game devices are all able to produce colour, and are commonly used to display colour-encoded information vi-

[1] A D65 light source is a light source that emits a spectrum of light that is similar to the distribution of light emitted by a blackbody at 6500 degrees Kelvin. D65 light corresponds roughly to a mid-day sun in western Europe and northern Europe.

sualizations. For example, computer games use colours to highlight team players, often using red for one team and green for another - a complication for the most common form of colour blindness [2]. As another example, the Apple iPod uses a blue highlight on a light background to indicate current selections (see Figure 1.2).

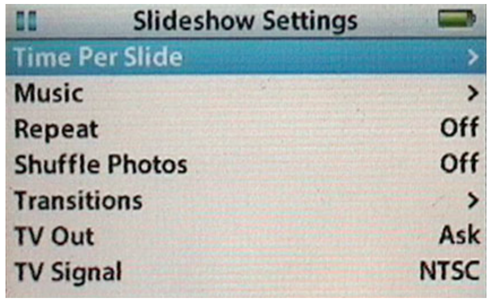

Figure 1.2: Highlighting colour on the Apple iPod. From: `http://www.techcrunch.com`

Colour blindness is usually genetic in origin and affects approximately 1 in 12 caucasian males and 1 in 200 caucasian females [5, 10]. When these individuals depend on information visualizations, the use of colour to encode information can cause problems. Information visualization often uses colour to identify categories of data, to highlight specific pieces of data, and to illustrate changes in continuous data. When a colour blind individual cannot distinguish between two categories of data, identify specific pieces of data, or recognize changes in continuous data, they will make errors in their interpretation of the visualization. These errors lead to frustration, increased response times, and reduced satisfaction. This can lead to lost productivity, increased corporate expenses, and in situations where colour is used to communicate safety-critical information, decreased public and personal safety.

1.3 Solution

The solution presented in this thesis is to construct a model of colour differentiation that is sensitive to the specific context of the colour differentiation task. Colour

[2]The most common forms of colour blindness are collectively known as *red-green colour blindness*, as the individuals with these forms of colour blindness have difficulty distinguishing between colours in the red-orange-yellow-green end of the spectrum.

differentiation is the process of discriminating between colours. Individuals who have colour blindness, or are under the influence of any other factors that influence colour perception, will have a reduced ability to discriminate between colours, and therefore have reduced colour differentiation abilities. Although the currently-used model has been referred to as a model of *colour perception*, it is used within this context of *colour differentiation*; therefore, modeling colour differentiation is the goal of this solution.

To make the presented solution sensitive to the specific context of a colour differentiation task, a series of measurements of the person's colour differentiation abilities are made. These measurements serve as calibration information that tunes a general model of colour differentiation to a particular individual in a particular environment. An example of a colour differentiation measurement is identifying the point at which an individual can distinguish between two colours. As these measurements are taken within the context of the colour differentiation task, all of the factors that influence colour perception are automatically incorporated into the model. By automatically incorporating colour perception factors into the colour model, the accuracy of the predictions about colour differentiation made by this colour model should be improved over the existing approach.

In addition to improved accuracy, there are two things that should be considered in the design of a colour differentiation model:

- Effort required to calibrate for an individual user. Calibration should use the minimum number of measurements that can reliably produce an accurate model. The process of taking a measurement should require a minimized effort.

- Over-specificity of the model. Over-specificity will result in the model having to be re-calibrated whenever the colour differentiation context changes slightly. The frequency of calibrations should be minimized.

1.4 Steps in Solution

Seven main steps have been carried out in this research: 1) the classification of factors that lead to atypical colour perception, 2) the identification of how colour is used in information visualization, 3) the definition of a general model of colour differentiation, 4) the construction of a calibration process that tunes this general model of colour differentiation to a particular individual and environment (a *context*), 5) development of a testing system, 6) evaluation of the assumption of linearity that was used to simplify the calibration process, and, 7) evaluation of the accuracy of the tuned model of colour differentiation.

1. *Classification of factors that lead to atypical colour perception.* In this step, external factors and internal factors that influence colour perception were identified and organized. External factors include such issues as the quality of the source of light, the presence of filtering material, and reflectance off of nearby surfaces. Internal factors include such issues as the presence of genetically-caused colour blindness, cataracts or yellowing of the lens with age, and cognitive disorder.

2. *Identification of how colour is used in information visualization.* Colour is commonly used to enhance information visualizations. This enhancement can come from using colour as a label, colour as a value, and colour to imitate reality.

3. *Definition of a general model of colour differentiation.* This step started with developing a working definition for colour differentiation. This definition was then used to construct a general model of colour differentiation. This general model contains differentiability values for every colour that can be produced in a digital environment, but requires over 100,000,000 measurement to achieve this. Reducing the number of measurements was the purpose of the fourth step.

4. *Determination of the calibration procedure for the general model.* The procedure developed in this step allows the general model of colour perception to be tuned to a specific context through the use of 48 measurements. The calibration process involves four stages: 1) testing users for basic colour differentiation abilities, 2) measuring users' differentiation abilities for 48 fundamental points in the general colour differentiation model, 3) extrapolation of measurements to ranges expected by the predictor constructor, and 4) construction of the colour differentiation predictor.

5. *Development of a testing system.* Java applications were constructed for the calibration procedure, the encapsulation of the model inside a predictor object, and the execution of the following two evaluations.

6. *Evaluation of the assumption of linearity.* This evaluation was performed to assess the feasibility of the assumption of linearity of the relationship between a colour and the ability to differentiate other colours from it. This assumption was used to implement the calibration procedure for the general model of colour differentiation.

7. *Evaluation of the accuracy of the tuned model of colour differentiation.* This study was performed to establish the accuracy of the calibrated model of colour differentiation. The accuracy was established by comparing the performance of the model to the performance of human participants for the same trials.

1.5 Evaluation

Evaluation of this research was performed using two user studies. The goals of these two studies were:

1. *Determine the feasibility of the linear assumption.* To achieve the reduction from over 100,000,000 measurements to 48 fundamental measurements in the calibration process, a simplifying assumption was made. This assumption was that the relationship between a colour and its differentiability from other

colours can be described using a linear function. The first study gathered differentiability data from six participants and found that a linear function could be used to describe differentiability relationships.

2. *Determine the accuracy of the context-specific model of colour differentiation.* The ultimate result of this research is a system (the predictor) that predicts whether two colours are differentiable or not. This predictor encapsulates the calibrated context-sensitive model of colour differentiation. The second study compared the performance of the predictor to the performance of eight human participants in determining whether two colours were differentiable or not. The results of this study indicated that the predictor agreed with the human participants in 86.3% of the trials.

1.6 Contributions

The primary contribution of this thesis is the provision of a colour differentiation model that is sensitive to the context of a colour differentiation task. This model is encapsulated in software that makes predictions about the differentiability of colours. This predictor has been shown to accurately model human colour differentiation abilities through a user study. There are also several secondary contributions: 1) identification of several factors that influence colour perception and differentiation, 2) a summary of colour use techniques in information visualization, 3) the development of a general model of colour differentiation in digital environments, and 4) a calibration procedure that transforms the general model of colour differentiation to a context-specific colour differentiation model.

1.7 Thesis Outline

The rest of the thesis is arranged as follows:

- Chapter 2 introduces background material related to this project. This includes placing this research in the larger context of human-computer interaction, an overview of colour perception, including many factors that influence colour perception resulting in atypical colour perception, and a description of existing applications of models of colour differentiation.

- Chapter 3 presents a summary of colour use techniques in information visualization, and how atypical colour perception disrupts these uses of colour.

- Chapter 4 gives a working definition of colour differentiation and describes the general model of colour differentiation, including some preliminary work that suggests the viability of the linear relationship assumption.

- Chapter 5 illustrates the process of minimizing the number of measurements required for the calibration procedure. The calibration procedure itself is also given.

- Chapter 6 presents a study that investigates the feasibility of the linear relationship assumption.

- Chapter 7 reports on a study that evaluates the accuracy of the calibrated model of colour differentiation.

- Chapter 8 gives the analyses and discusses the implications of the results from the Chapter 7 study. This chapter also discusses some of the design decisions involved in this project.

- Chapter 9 presents a summary of the research, the overall contributions of the thesis, and a set of topics for future work.

Chapter 2

Background

This research is based on eight areas of previous work: consideration of non-typical users in human-computer interaction, non-typical users with colour vision disabilities, fundamental colour concepts, factors that influence colour perception, manifestations of atypical colour perception, colour vision tests, current models of colour differentiation, and existing applications that use models of colour differentiation.

2.1 Human-Computer Interaction

The computer science discipline of human-computer interaction (HCI) is focussed on the interface that exists between digital devices and their human operators. As defined by the Association of Computing Machinery [25], a working definition of HCI is the following:

> Human-computer interaction is a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them.

With this focus on interactive systems for human use, researchers in this discipline often construct models of human users based on measurements of their physical properties (e.g., height), or their abilities (e.g., reach when seated) [56]. As a consequence of this modeling technique, users are often lumped into large categories, such as 'disabled' or 'novice'. Alan Newell discusses the negative results of this lumping process in [45], where he makes a case for extra-ordinary human-computer interaction. In his discussion, Newell identifies that "...every human being has a set of abilities and characteristics, some of which can be described as 'ordinary' and some of which are very obviously 'extra-ordinary.'

Any computing environment where the design has been based on some notion of an 'average user' potentially alienates any individual that does not or can not conform to the average model because of the presence of these extra-ordinary abilities or characteristics. These individuals can be thought of as non-typical users.

In HCI, those devices which address the needs of non-typical users are collectively known as *assistive technology* or *assistive devices*. The general approach to developing assistive devices is to identify a non-typical user group, identify the problems experienced by this group, and then develop technological solutions to these problems [63, 16].

2.2 Non-Typical Users

The benefit of modeling for a single type of user brings with it the cost of alienating non-typical users. Many potential users of computer systems could be classified as non-typical users. The elderly present a growing population which has historically been neglected by HCI [8, 15, 1, 44, 11]. Differentiating between genders in software development and use [9] has generally been downplayed. Children present non-standard abilities that interface experts should consider [7], and consideration of cultural variation for globally-used computer systems [39] is essential in a shrinking world. More 'traditional' forms of non-typical users are those with cognitive impairment [46], physical disabilities [55], and perceptual impairments [33]. Any individual in a non-typical environment is also a non-typical user. This was explored early by Newell [45], and continues to be developed by Sears [3], with the notion of situationally-induced impairments and disabilities.

In this thesis, those individuals with atypical vision will be examined. Although there are many forms of vision-related non-typical users (e.g., low vision [19]), the specific non-typical users discussed in this thesis are those that have non-typical or atypical colour perception abilities. As will be shown, atypical colour perception can be caused both by traditional notions of disabilities (such as colour blindness or yellowing of the lens with age), as well as the notion of situationally-induced colour vision impairments (such as poor environmental lighting or the presence of filters). These are identified as *internal* and *external* factors that influence colour perception.

2.3 Fundamental Colour Concepts

In this section, a brief introduction to colour theory is given. Fundamental concepts of colour in the natural world, how the human eye processes incoming light in order to detect colour, and tristimulus colour models used for the digital representation of colour are discussed. Colour terminology used throughout this thesis is also defined.

2.3.1 What is Colour?

Colour, like other perceived qualities of our environment, is a property that does not exist beyond our detection of it (i.e., it is not a property of the world, but resides inside our heads). Just as sound only exists as longitudinal waves in air until the waves contact our auditory sensors (ears) and produce some neurological response, colour only exists as waves until the waves enter our visual sensors (eyes), and produce a neurological response. In nature, colour results from the incidence of a visible spectrum distribution upon the retina of the eye. This spectrum is composed of varying amounts of electromagnetic radiation (EMR) at wavelengths between 380 and 780 nanometers (see Figure 2.1). EMR within this range is generally called *visible light*, or simply *light*.

Although digital colour output devices (such as monitors, projectors, and printers), often produce colours similar to nature, they do not produce a full spectrum distribution as is produced with natural colour. The reason artificial colour sources can approximate nature so well is directly related to the manner in which the human eye detects light and colour.

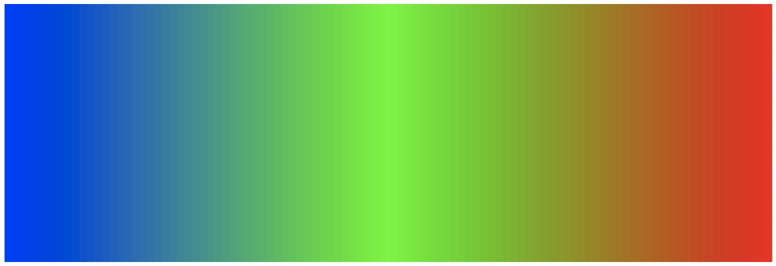

Figure 2.1: An approximate full spectral distribution. Note that this image progresses from short wavelength (high frequency) on the left to long wavelength (low frequency) on the right.

2.3.2 Colour Detection

The retina in the typical human eye contains three categories of colour sensitive photoreceptors called *cones*.[1] Each type of cone can be modeled by a function that indicates the degree of cone stimulation achieved by energizing it with a given pure frequency of light (see Figure 2.2). It can be seen that each cone has a unique response to each possible frequency of light. Using a normalized cone response such as this, it can be determined how each type of cone will respond to a given spectrum distribution by multiplying the normalized cone responses with the spectrum's normalized frequency distribution [59]. The degree of stimulation of each cone, and the ratios of stimulation between cones, provide the initial signals to the human visual processing system that lead to the perception of colour. This system of three colour-sensitive photosensors is called the trichromatic theory of colour vision, also known as the Young-Helmholtz theory, named after its early proponents [18].

The brain contains of a number of processing centers that aid the discernment of shapes, colours, edges, etc. One of the earliest of these visual processing centers is the lateral geniculate nucleus (LGN). This area of the brain is where the optic nerves connect to the brain. In this center, some basic processing of the optic nerve signals is performed. This processing leads to the perceptual opposition of colours such as red and green, and yellow and blue. In addition to these oppositions, this center is responsible for the dominant role brightness plays in the visual system. As

[1] Cones are effective in well-lit environments (such as outdoors). The eye also contains low-light photosensors called *rods*. Rods do not generally contribute to colour detection.

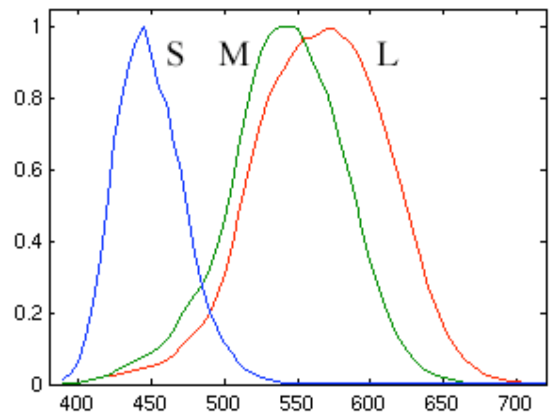

Figure 2.2: Frequency response functions for the three types of cones. The blue function line is for the **S**hort wavelength sensitive cones, the green function line is for the **M**edium wavelength sensitive cones, and the red function line is for the **L**ong wavelength sensitive cones. Wavelength is the independent variable (x-axis) and the normalized cone response is the dependent variable (y-axis). From: `http://en.wikipedia.org/wiki/Cone_cell`

an example of this dominant role, Figure 2.3 contains two copies of the same image, one in colour and one in greyscale. It can be seen that although colour adds some additional information, the brightness of the image carries the majority of visual information.

Introduced in 1872 by Ewald Hering [24], the opponent process theory of colour is based on the notion that humans tend to perceive four distinct ‘primary’ colours - red, yellow, green, and blue. Although considered to be at odds with trichromatic colour theory for some time, it is now recognized that both systems are at work in the human visual processing system. The cones provide initial values to the LGN which processes the cone information to determine values for three new channels, a red-green channel, a yellow-blue channel, and an overall luminance channel. ‘Opponent’ comes from the opposition between red and green, as well as yellow and blue, such that reddish-green and blueish-yellow are generally unimaginable hues. In opponent process colour theory, each opposition pair is encoded as a single channel, where cone-level detection of one of the colours stimulates the channel, and the opposing colour inhibits it. This is illustrated in Figure 2.4, in which the red-green channel

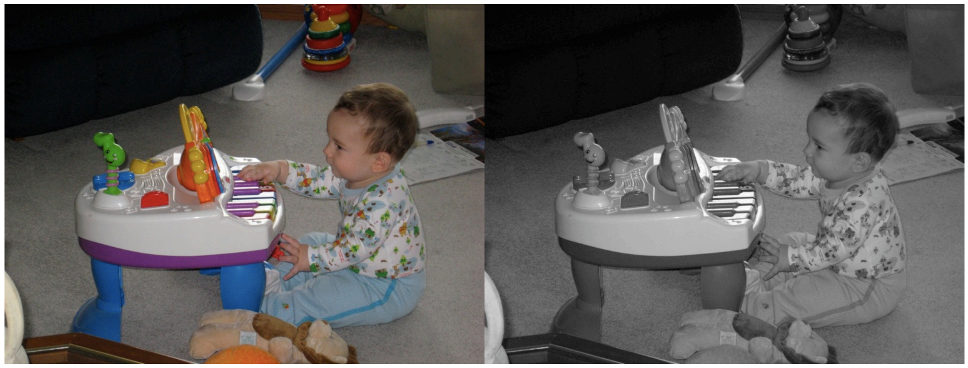

Figure 2.3: The image on the left is in colour with some bright colours present. The image on the right is the same image with all colours removed and only greyscale (luminance) present. As can be seen, the colour of the first image adds little additional information in comparison to the amount of information luminance presents.

is encoded using information from the long and medium wavelength cones (long - medium), the blue-yellow channel is encoded using information from all three cones ((long + medium) - short), and the luminance/brightness channel is encoded using information from the long and medium wavelength cones (long + medium). Note that the short wavelength cones contribute nothing (or very little) to the overall perception of brightness. This is supported by the standard colour to greyscale conversion equation of $grey = (0.3 \times R) + (0.59 \times G) + (0.11 \times B)$, in which the blue channel (which mostly stimulates the short wavelength cone) contributes only 11% to the total intensity for a grey value.

Generally, different incident spectrums stimulate the cones differently, and hence are perceived as different colours. Occasionally, two different frequency distributions will stimulate the cones identically, thereby being perceived as two identical colours, even though their respective spectra are different. Different frequency distributions that produce the same colour response are called *metamers*. This principle of *metamerism* is the basis for digital colour representation.

2.3.3 Digital Colour Representation

Metamerism introduces an interesting possibility. If each cone type could be stimulated independently of the remaining cones, then *every* possible colour could be

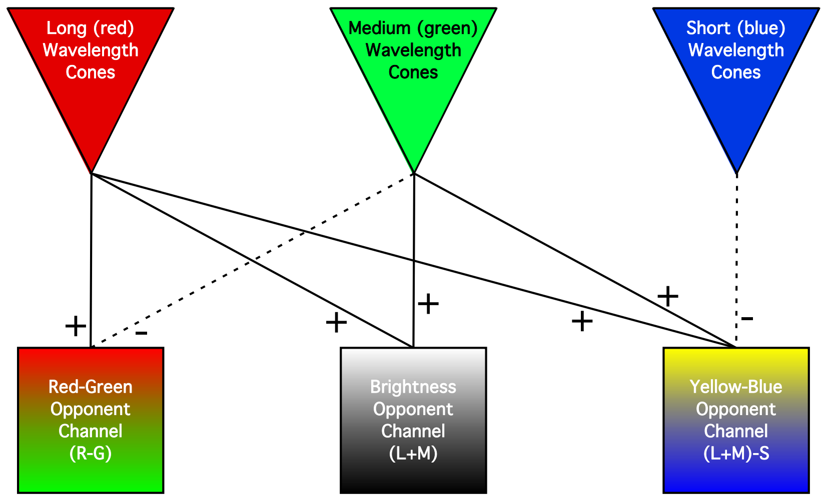

Figure 2.4: Opponent process colour model. (Long-medium) gives red-green opponent channel, ((long+medium)-short) gives yellow-blue opponent channel, and (red+green) gives the white-black or brightness opponent channel.

produced by sufficiently stimulating each of the cone types in this manner. Unfortunately, it is not possible to independently stimulate cones using light, because the cone sensitivities overlap. Independent stimulation, however, can be almost achieved by carefully selecting a small number of lights that emit a particular wavelength. If one light is chosen to mostly stimulate short-wavelength cones, another light to mostly stimulate medium-wavelength cones, and yet another light to mostly stimulate long-wavelength cones, then it will be possible to produce many different colours by varying the intensities of these three lights.

Using three lights (called *source lights* or *primary lights*) to produce a wide range of colours is the core concept behind digital colour representation. CRT monitors contain three types of phosphors that emit distinct frequencies of light when stimulated by a stream of electrons. These three phosphors produce red, green and blue light, respectively. The frequencies of these three colours are sufficiently distinct to allow controlled stimulation of the long, medium, and short wavelength cones of the human eye. To generate a given colour, the red, green, and blue phosphors are stimulated to emit various proportions of red, green, and blue light. This light enters the eye, striking the retina, and stimulating the long, medium, and short wavelength sensitive cones. The stimulation of these cones triggers a sensation of colour. This is the basis of the Red Green Blue (RGB) colour model used in all current computer graphics systems. This model has been extended to colour projectors, LCD displays, and plasma televisions.

RGB is an example of an *additive* colour model [23]. In the example of a CRT monitor, no stimulation of any of the phosphors produces black, and colours grow in intensity as the stimulation of the phosphors is increased, finally resulting in white when all phosphors are maximally stimulated. A *subtractive* colour model operates on the opposite principles, where a total lack of channel 'stimulation' produces white, and the total 'stimulation' of the colour channels produces black [20]. An example of a subtractive colour system is an artist's use of paint. The paper or canvas is (usually) white, and if no paint is applied, white is produced. If the painter mixes and applies all of the colours in her palette, then a colour approaching black will be produced. Why a pure black is not produced (and many more details of subtractive colour models) will not be further discussed here.

RGB is not a very natural representation of colour to work in, as individuals often think in terms of subtractive colour models, rather than additive colour models. Subtractive models of colour provide concepts of *tints* and *shades*, where a tint of a colour is obtained by adding white to it (thereby lightening the colour), and a shade of a colour is obtained by adding black to it (thereby darkening the colour). Separate *tones* of a colour can be obtained by mixing equal amounts of black and white (grey) with the colour. This terminology is used occasionally in this thesis.

The concept of lightening and darkening a colour is so pervasive (and useful) when discussing colour, that conversions from RGB to more *perceptual* models of colour have been developed. These are often based on concepts of *hue*, *saturation*, and *value*. Hue is the specific colour, such as red, green, yellow, or blue. Saturation is a measure of the vividness of the hue. Value is the brightness of the hue, with minimum brightness being black. Any hue with no saturation and maximum brightness is white. The original concept of HSV as well as algorithms for converting between RGB and HSV are described in [57]. See Figure 2.5 for a visual representation of

HSV space.

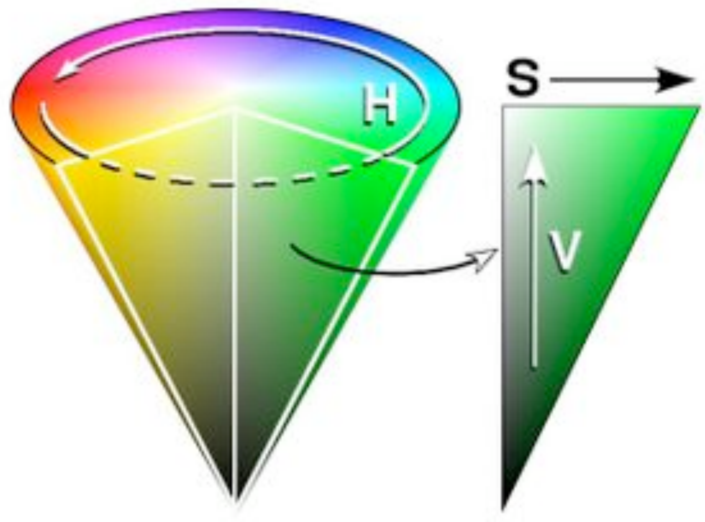

Figure 2.5: HSV colour space represented as a cone. Value (lightness) follows a central line from bottom to top, saturation is the radial distance from this line, and hue is the angle of the saturation line from some arbitrary baseline (usually red). From: `http://en.wikipedia.org/?title=Hue`

2.4 Factors that Influence Colour Perception

The world in which humans reside influences how colour is perceived. On a large scale, these influencing factors can be categorized into two classes: external factors and internal factors. External factors relate to the environment's influence on colour perception by producing and modifying the spectral distribution of visible light incident on the perceiver's eyes. Internal factors include the physiological variations from person to person that alter how the individual perceives colour.

As outlined in Section 2.3.2, human colour perception is a combination of external and internal systems. External systems involve sources of light and how these interact with an observer's surroundings to produce spectral distributions that enter the eye. Internal systems include the eye and its preprocessing abilities (including transforming the incident spectral distribution into a neurological response), as well as the optic nerve and numerous visual processing regions in the brain. The external and internal factors that change the 'normal' model of human colour perception will now be examined in more detail.

2.4.1 External Factors in Colour Perception

Before visible light enters the eye to be perceived as colour, it is produced and then modified by the environment surrounding the person doing the perceiving. This section examines factors influencing the sources of light and the factors that potentially modify the light before it enters the eye.

Light Sources

A source of light is necessary for colour perception. For the purposes of colour reproduction, the most accurate sources of light are those that mimic the wavelength distribution of the sun on a clear, sunny day. This distribution is almost uniform in that the intensity of any given visible wavelength is roughly the same as any other visible wavelength. An ideal uniform light source wavelength distribution is shown in Figure 2.6. An object has a perceptible colour because some of the source light is reflected off the object's surface, and some of it is absorbed (some is also scattered). Under a variety of source lights, if each has a uniform light distribution, the absorption and reflection remains the same, and the object is perceived as having the same colour, regardless of which uniform source light is illuminating the object [67].

It is often assumed that the frequency distribution of commonly used light sources approaches a uniform distribution, but this is not the case. Incandescent sources of light have a semi-uniform wavelength distribution that is biased towards the red end of the spectrum (causing the light to have a yellowish colour). Fluorescent lights have an even more red-biased distribution with little coverage of the blue-violet end of the spectrum, save for some narrow peaks. These peaks result from the types of phosphors that coat the inside of the fluorescent light. These peaks result in fluorescent lights appearing to produce an even white light (to the human observer), when in fact these lights do not produce a uniform distribution of light [54].

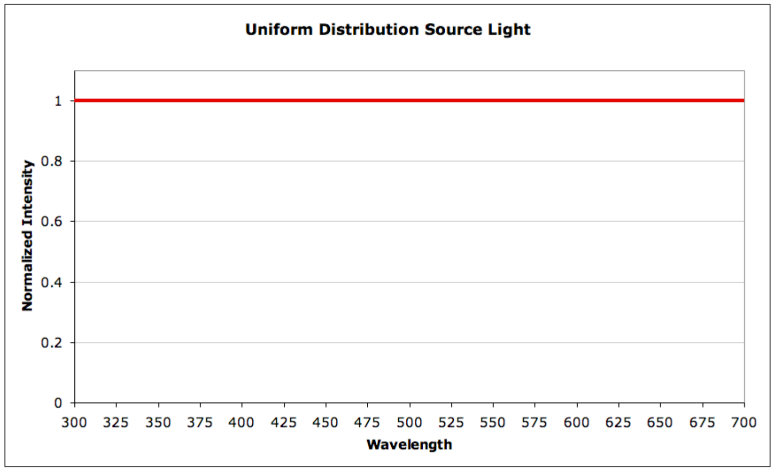

Figure 2.6: A hypothetical and perfect uniform distribution light source. A real light source will have some variation between wavelength intensities, but if it contains relatively strong intensities for all visible wavelengths, then it is considered uniform.

Variations between light sources sometimes cause the same object to have a different colour, depending on the light source [26]. This can be illustrated with an example. Imagine some object has a surface reflectance that is high at the violet end of the spectrum and drops off steadily as the wavelength lengthens. In sunlight, this object should appear to be a tint of violet-blue. In incandescent light, this object should appear a more neutral light grey colour. In fluorescent light, this object should appear a shade of purple, as the fluorescent light produces strong red light, as well as peaks in the blue and violet range. Reflection off of this hypothetical object will give strong incident reds and blues, with weaker middle wavelengths, giving the object a muted purple colour.

When viewing an object the source of light can have a considerable influence on the colour that is perceived for the object [67]. This may seem strange to many people, because independent of the source of light, most objects remain the same colour - e.g., a red apple is red in the morning sunrise and in the afternoon sun, under fluorescent and incandescent lights, and even in highly coloured light. The reason for this is the concept of *colour constancy*. Colour constancy results from higher-level cognitive processing that causes the distribution of source light to be considered when the colour of an object is examined. Essentially, if the source light strongly emits green light, and weakly emits red and blue light, the brain uses the signals from the cones to determine that this is the case, and therefore adjusts the colours to reduce the effect of this biased light. This is also evident when examining the same colour in shade and in sunlight. The colour will generally appear to be exactly the same. Colour constancy can be used to create interesting illusions, as shown in Figure 2.7, in which the squares labeled 'A' and 'B' are exactly the same colour, even though they do not appear to be so.

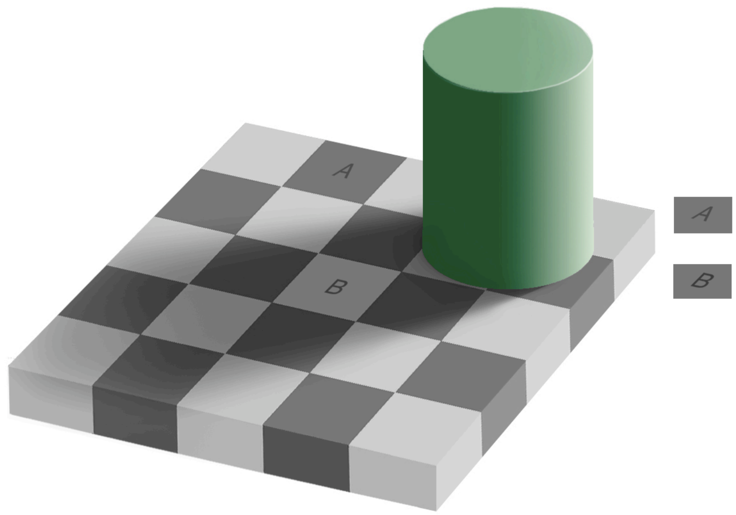

Figure 2.7: A common optical illusion that uses colour constancy to trick the viewer. The squares labeled 'A' and 'B' are identical shades of grey. From: `http://en.wikipedia.org/wiki/Color_constancy`

Computer monitors, televisions, and digital projectors do not depend on an external light source to produce colour (because they are additive colour devices and hence have their own source of light), but vary widely between devices in the distribution of light they produce. Considerations for colour production in digital devices include the software and operating system generating the original values for the RGB colours (including any colour profiles present as well as gamma settings), the hardware that produces the raw image for the display device (generally a graphics card of some sort), and the display itself. Operating systems each take their own approach to colour, and each software package may use colour in very different ways. Graphics cards are produced by different manufacturers and have varying capabilities and quality. Displays vary greatly from manufacturer to manufacturer and have been changing in quality over the past several years as the use of LCD monitors has increased.

Another consideration of display hardware is the brightness of light that is hitting the screen. If a machine (such as a laptop or handheld computer) is to be used outside, sunlight will be incident upon the screen. This will probably wash out the display. The amount of light produced by a monitor has a maximum limit imposed by the light-producing hardware. In environments where there is a bright light source (e.g., outside on a bright day), our eyes automatically adapt to restrict the amount of light entering the eye. As the maximum brightness of a monitor is limited, less light from the monitor will enter the eye in such a situation. This results in the colours produced by the monitor being 'washed out' or desaturated as well as having a reduced perceived brightness. Both of these factors will influence how the colours produced by the monitor are perceived.

Another external factor which may influence colour perception is the display hardware itself. The colours selected by the designer of an information visualization may not be accurately reproducible on the viewer's display. This may be a result of poor monitor calibration or manufacturing processes, or errors in the display adapter hardware or software.

Light Modifiers

Light that has been emitted from a light source can be modified in two main ways before it reaches the eye. The light may be modified by reflectance off of any surface or it may pass through a material that acts as a filter [59].

When the environmental light sources strike a surface, some of the light source frequency distribution is absorbed by the surface, and the remainder is reflected. This reflected light enters the eye and triggers the perception of colour. Of course, the reflected light may reflect off other surfaces (with different light absorption characteristics) before it enters the eye. Reflection and re-reflection (and re-re-reflection, etc.) influence the distribution of light that enters the eye, and thereby influences colour perception. As an example, consider examining the colour of an object in a room that is illuminated by indirect lighting. The source lights emit their light, which is reflected of the walls of the room before it is reflected off the object one is examining, and into the eye. If the colour of the walls absorbs certain spectra, then the light incident on the object will be significantly different than if the light came directly from the source lights, and the perceived colour of the object may be different. Colour constancy does help to negate some of these issues with reflectance. Of course, this only affects colour which is reflected (such as in print), but not colours that are the product of coloured light sources (such as computer monitors).

Source light generally travels through the air, reflects off an observed object, and then enters the eye of the observer. Air is generally clear (it transmits all visible light equally), and therefore does not influence the colours perceived. In some circumstances, the air can contain particulates (such as smoke or fog) which act as filters, selectively absorbing some wavelengths, while allowing others to pass through uninhibited. Many other potential filters exist, including optics such as glasses and contact lenses, sunglasses, and windows. This can influence colour in digital environments, since the display screen material may also act as a filter.

2.4.2 Internal Factors in Colour Perception

Internal factors that influence colour perception include all the physiological variations from person to person. These include genetically-linked variations, such as colour blindness, as well as factors related to physical or mental abnormalities (such as disease) in an individual that alter the way in which the individual perceives colour.

Approximately 8% of men and 0.5% of women have a genetic condition which causes atypical colour perception [10]. This is commonly called *colour blindness*, or *colour vision deficiency*. Each variety of colour blindness is now described, as well as how severely the colour perception of individuals with such a condition is affected. Genetic and non-genetic causes of colour blindness are discussed.

Genetic - Anomalous Trichromacy

Figure 2.2 illustrates the sensitivity distribution by cone type in typical colour perception. The wavelength at which a maximal cone response occurs is called the *peak sensitivity* for that cone type. In *anomalous trichromacy*, the peak sensitivity of a cone type is shifted. Long wavelength sensitive cones are shifted towards the medium sensitive cones peak sensitivity (left), and medium wavelength sensitive cones are shifted towards the long wavelength sensitive cones peak sensitivity (right). Anomalous trichromacy with short wavelength sensitive cones is a rare and poorly understood condition, but is probably similar to its short and medium wavelength counterparts [37]. Long wavelength cone shifting is formally called *protanomaly* (where 'prot' is a reference to the 'first' cone type), medium wavelength cone shifting is called *deuteranomaly* ('deut' refers to the 'second' cone type), and short wavelength cone shifting is called *tritanomaly* ('trit' refers to the 'third' cone type). There are also rare cases of individuals who have both protanomaly and deuteranomaly. This condition is called *extreme anomalous trichromacy* and is outlined in [6].

Shifting in the peak sensitivity is governed by complicated genetics (see [43] for a detailed explanation). Although it will not be discussed here, let it suffice to say that the amount of shift in peak sensitivity is not consistent between anomalous trichromats. The central result of this variety in sensitivity shift is that the degree of colour blindness varies significantly between anomalous trichromats, from no noticeable effects, to nearly complete loss of fidelity on the shifted cone's respective data channel. A detailed explanation of the effects of this shifting is given in Section 2.5.

Genetic - Dichromacy

Individuals with *dichromacy* are missing one of the types of cones. In this situation, either the long wavelength, medium wavelength, or short wavelength cones are completely missing. If the long wavelength cones are missing, the person is diagnosed as being *protanopic*. Likewise, *deuteranopic* and *tritanopic* are used to describe the condition of missing the medium or short wavelength cones, respectively.

Unlike anomalous trichromatism, there is little variation in the severity of colour blindness experienced by dichromats of a given type. As such, all protanopes (deuteranopes, tritanopes) generally experience similar colour perception.

Genetic - Monochromacy

The concept of dichromacy can be extended to the condition in which an individual is missing two types of cones; this is called *monochromacy*. The single remaining cone (often the short wavelength cone) is able to detect variations in lightness, but no colour information is passed along to visual processing areas of the brain. As such, monochromats have essentially no colour perception beyond shades of grey (except in certain circumstances where the *rods* contribute colour-useful data in low-light conditions [5]). These individuals often will have slightly reduced visual acuity in addition to the lack of colour perception.

Genetic - Achromatopsia

The final established type of genetically-determined atypical colour perception is the situation in which no cones are present. This is often called *achromatopsia*, but this name is also used to describe monochromacy. As only rods are present, these individuals have poor visual acuity, are extremely sensitive to bright light, and have no colour perception beyond the greyscale vision of monochromacy [53].

Genetic - Tetrachromatism

Genetic colour blindness issues related to the long and medium wavelength sensitive cones are carried on the X-chromosome. As men have one X-chromosome, colour blindness is expressed more frequently in men. Women, having two X-chromosomes, need to have both containing colour-blind genes in order for the woman to be colour blind, as the presence of a single colour-blind chromosome would be masked by the other non-colour-blind chromosome.

It has been hypothesized that occasionally this masking may not occur properly. This results in a fourth cone type to be expressed in the woman's retinas. This fourth type is believed to have a peak sensitivity somewhere between the long and medium wavelength sensitive cones, and should allow greater colour distinction than typical trichromat colour vision. Many questions remain about this hypothesis.[2]

Non-Genetic Factors of Colour Perception

As the human visual system is a neurological one, any chemical or physical stimuli that effect the central nervous system can modify how an individual perceives colour. As such, it has been found that many anti-depressant prescription drugs and other medications (such as Viagra) can influence colour perception (see [5, 59]).

As the eye is an organ like any other component of the body, it can be damaged through exposure to certain chemicals or environments. In many cases, the retina in particular will be negatively impacted. As short wavelength cones only comprise about 10% of the retina's colour sensitive photosensors, these are much more susceptible to retinal damage. When the short wavelength sensitive cones are destroyed, the resulting atypical colour perception is very similar to tritanopic colour blindness. Exposure to certain organic compounds (such as styrene) can result in this condition, as can long-term diabetes [10]. Any condition which leads to retinopathy (damage or death of the retina) will effect colour perception negatively. Tritanopic type effects can also accompany yellowing of the lens with age, as well as clouding of the lens through the onset of cataracts [5].

2.5 Manifestations of Atypical Colour Perception

In Section 2.4, it was discussed how the typical model of colour perception covered in Section 2.3.2 can be modified by various genetic, internal, and external factors, resulting in atypical colour perception. In this section, the manifestation of these colour perception difficulties are explored further.

[2] For example, how is the additional cone distributed in the retina? How can an accurate test for the condition be developed (by trichromat researchers)? Does the optic nerve have the necessary bandwidth to transfer an additional channel of colour information? How are the visual processing centers of the brain modified to accommodate the new information? How does the opponent process colour theory (see Section 2.3.2) handle the additional channel? [38, 43]

2.5.1 Red-Green Colour Blindness

Opponent colour theory gives a useful model for understanding why the prevalent forms of colour blindness (protanopic and deuteranopic anomalous trichromatism and dichromatism) are commonly referred to as *red-green colour blindness*. The long and medium wavelength sensitive cones both contribute to the red-green opponent channel, one by stimulating and one by inhibiting. To demonstrate the manifestation of genetically-determined atypical colour perception, imagine the following idealized situation:

> Assume that each cone type can be stimulated in complete isolation. If a long wavelength cone at maximum stimulation produces ten 'points' of stimulation (+10) to the red-green opponent channel, and the medium wavelength cone at maximum stimulation produces ten 'points' of inhibition (-10), then the individual can experience a range of twenty values (perhaps unique colours) on the red-green opponent channel. Now assume that the colour blind individual's long wavelength sensitive cone has been shifted toward the medium wavelength cone's peak sensitivity. This will nullify the stimulation in complete isolation premise above. Perhaps when the long wavelength cone is stimulated, the medium wavelength cone is stimulated at 30% of the long wavelength cone's stimulation (and vice versa). Hence, when the long wavelength cone is maximally stimulated (+10 points), the medium wavelength cone is stimulated as well (-3 points). Likewise when the medium wavelength cone is maximally stimulated (-10 points), the long wavelength cone will be partially stimulated (+3 points). This results in a reduced range of fourteen values. Dichromats are simply missing one of the cone types, giving a further reduced range of either 0 – 10 points for protanopes, or -10 – 0 points for deuteranopes, according to the above assumptions. As a result, the number of 'steps' between red and green hues is diminished, thereby allowing these completely 'opposite' colours to be confused. This model can be extended to tritanopic colour blindness as well.

2.5.2 Trichromatic Manifestations of Atypical Colour Perception

To further understand the manifestations of atypical colour perception, we can use the tristimulus colour models presented in Section 2.3.3. Imagine a colour model in which the amount of stimulation of each separate cone type is represented on a separate linear scale, say between 0.0 and 1.0. This would allow another tristimulus colour model to be developed in which a separate dimension is given to each cone type. This three-dimensional volume models every colour that can be seen by an individual. Figure 2.8 contains an approximation of this volume. Each point in the space represents a unique stimulation pattern for each cone, which will be perceived as a unique colour.

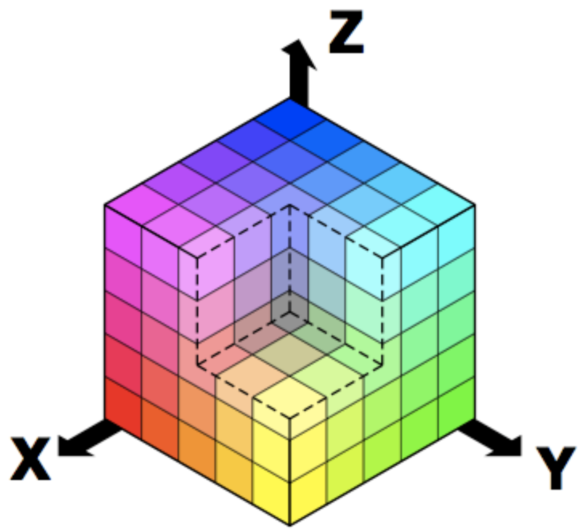

Figure 2.8: A three-dimensional representation of separate cone stimulation. Note that the colours shown are from CIEXYZ colour space. An approximate mapping represents the long wavelength sensitive cone on the X dimension, the medium wavelength sensitive cone on the Y, and the short wavelength sensitive cone on the Z. From: `http://en.wikipedia.org/wiki/Color_models`

This model allows a clearer understanding of the atypical colour perception caused by missing cones. Dichromats will have one of the dimensions of the cube removed (the dimension corresponding to the missing cone type). In effect, this reduces all colours along a line parallel to this dimensional axis to collapse to a single perceived colour. For example, if the long wavelength cone is missing (protanope), and it is assumed that this cone only detects red, then it can be seen that all colours that only differ in the amount of red contained in them would collapse to an identical point in the now two-dimensional colour space. This effectively results in the protanopic individual confusing all colours that only differ in their respective amounts of red.

As monochromats retain the use of one type of cone, two dimensions of the colour volume are compressed to a single line (one dimension). This reduces the number of colours perceptible to these individuals so significantly that only shades of grey are seen.

Anomalous trichromatism also results in a compression of the 'volume' of colours perceived, but in a less severe manner as dichromatism [59]. Although dichromats have an entire dimension of the colour volume missing, the volume is partially compressed in a non-linear fashion for anomalous trichromats, such that the number of colours is reduced, but not as severely. This compression occurs along the axis corresponding to the cone that has its peak sensitivity shifted. The greater the shift, the greater the degree of compression. Using the volume presented in Figure 2.8, we can model typical colour perception as an overlapping subdivision of this colour space into spherical blobs. All colours inside a single blob constitute a 'single' perceived colour. In anomalous trichromatism, this sphere 'stretches' to become a cigar or hot dog shaped volume, along the shifted cone's axis, the extent of this stretching depending upon the severity of sensitivity shift. This extension of the shape causes more data points to be encapsulated in the same volume (inside which all colours are perceived identically). As a result, the total number of distinct colours is less in anomalous trichromatism than typical colour perception.

2.5.3 Darkening of the Red End of the Spectrum

Protanopic and protanomalous (long-wavelength cone) individuals have an additional constraint that influences the colours they perceive. Spectral distributions that contain solely long-wavelength light only trigger a response in the long-wavelength sensitive cones. Referring back to Figure 2.2, it can be seen that incident light of 675 nm will only activate the long wavelength sensitive cones. Now consider the shifted sensitivity of a protanomalous individual. Incident light at 675 nm will stimulate the long wavelength cones to a lesser degree than in typical colour perception. This manifests as a reduction in the lightness of red colours. More incident light is needed to stimulate the long wavelength sensitive cones sufficiently to cause the perception of red. Overall, individuals with long wavelength sensitive cone disorders tend to perceive red colours as less distinct and vivid, and can often confuse them with dark colours such as black, navy, dark brown and dark green.

2.6 Tests of Colour Vision

Many tests exist that evaluate the colour vision abilities of an individual [5]. According to [13], these tests can be divided into four categories:

1. Pseudoisochromatic plate tests.
2. Arrangement tests.
3. Matching tests.
4. Naming tests.

Pseudoisochromatic plate tests involve examining whether an individual is able to identify a number embedded in a background by differentiating it based on colour differences only. The most commonly used version of this test is Ishihara's Test [32]. This test is used to identify the presence of red-green types of colour blindness. It does not identify the severity of colour blindness, nor does it identify any short-wavelength-sensistive cone abnormalities.

Arrangement tests involve having the individual being tested sort a set of coloured tablets by hue. The most common forms of this test are the Farnsworth Munsell 100 Hue Test and its simplified 15 hue version [17]. A score is derived after the sorting is completed by comparing the subject's order to a predetermined order. This test cannot distinguish colour vision normals from mildly anomalous trichromats and does not always differentiate protanopes from deuteranopes.

Matching tests involve the subject manipulating one colour until it matches another colour, which does not change. In this test, the subject adjusts the amount of pure spectral red and green light to match a pure spectral yellow light. Once the colours are matched, the amount of red and green in the adjustable light indicates precisely the type and severity of colour blindness the subject has. There are similar extensions to provide blue-yellow colour blindness testing. The devices that perform these tests are called anomaloscopes [14]. These devices do suffer some difficulties due to learning affect with yellow-blue colour vision testing when retinopathy or yellowing of the lens is present.

Naming tests are occupation-specific tests which involve subjects identifying the colours of a set of coloured lanterns. Industries such as railway, maritime, and flight require the operators of their respective vehicles to be able to easily recognize the colour of coloured lights used to signal and identify. These tests do not attempt to identify either type or severity of colour blindness.

In addition to the traditional tests described above, attempts have been made to develop colour vision tests that run on computer systems using colour displays [22, 2, 50, 4]. These papers present some early success in detecting atypical colour perception, but could not achieve the precision such as that attainable in anomaloscope systems.

2.7 Current Models of Colour Differentiation

Current models of colour differentiation are based on an algorithm presented in [64]. To visually present how colour blind individuals (strictly speaking, just protanopes and deuteranopes) see the world, a sample image is converted from 'normal' colour to 'colour blind' colour using the following steps:

1. Using a pre-defined orthogonal transformation, translate the original image pixel colours into a colour representation that encodes colour as stimulation levels for the three types of cones.

2. Remove the long-wavelength cone (for protanopes) or medium-wavelength cone (for deuteranopes) stimulation values for every pixel in this representation.

3. Adjust the stimulation of the remaining cones to maintain the overall brightness of the image.

4. Using the inverse of the orthogonal transformation used in Step 1, translate the cone stimulation representation back to the original colour representation.

To extend this algorithm to make predictions about colour differentiability, one additional step is required:

5. Compare regions of colour in the modified image against the original image. If regions that are different colours in the original image are the same colour in the modified image, then these colours are considered not differentiable (confused).

Modeling colour differentiation in this manner requires two main components. The first component is the orthogonal transformation and its inverse. These transformations are generated using the spectral values for the RGB primaries of the monitor, the overall white balance of the monitor, and a general model of the spectral response curves for humans. The second component is knowledge of the type and severity of colour blindness under consideration. This information is potentially available from colour vision tests.

This approach does not work well. The spectral values for the RGB primaries of monitors are generally not known, and vary from device to device [40]. The overall white balance of the monitor may not be known and will vary with age of the monitor [12]. Both of these issues can be solved using frequent recalibration, but most computer users do not own a calibration device. It is also known that cone spectral response curves vary between humans, limiting the effectiveness of a general model [18]. As has been shown above in Section 2.6, only one type of test available (anomaloscopes) gives precise information about the type and severity of colour blindness to be modeled. Even though these devices are available, it seems that they are not commonly used. Although this is personal conjecture, it seems that only the detection of colour blindness is generally desired, and not information about the specifics of the diagnosis. As an example, I have known that I am colour blind from a very young age. Despite actively researching colour vision for the past three years, I am still unsure about the type of colour blindness I have. If I have had this much difficulty, what are the chances a non-expert will know these details?

Secondary reasons that make this modeling approach fail are the myriad of internal and external factors that influence colour perception (Section 2.4), leading to atypical colour perception. These factors move well beyond genetic colour blindness, and are not considered in the model described above at all.

2.8 Applications that Use the Current Models of Colour Differentiation

Systems that use the model described above in Section 2.7, are those that recolour an image such that the recoloured version is more useable by a particular colour blind individual. I will refer to these systems as colour adaptation tools (CATs).

SmartColor [66] allows the designer of a visualization to specify properties for colours used in the visualization. These properties then serve as constraints that steer a colour search algorithm that finds a colouring palette that is colour blind 'friendly' and agrees with the original properties specified by the designer. This system uses a model of colour differentiation to identify differentiable colours according to the specifications of the designer.

When navigating the web, many websites use colour schemes that result in confusion for colour blind users of the website. To address this issue, edge services [28, 29] can be deployed that automatically detect problem colour use and recolour the website to help eliminate the problem. These systems use a very basic model of colour blindness, assuming that red-green types of colour blindness can be assisted by increasing the colour contrast between those regions that are red and those regions that are green.

Images provide the basic encoding for colour. Videos, websites, visualizations,and desktop systems can all be thought of as using images to represent information. Colour adaptation tools that address images have been developed [34, 35, 31], including some that use modified greyscale transformation algorithms to modify the presentation of colour [48, 49].

The methods used to recolour an image focus mainly on the problem of selecting new colours. As there is a large number of choices for the new colours, genetic algorithms have been proposed as a means to identify new colours [30, 61], as well as systems that aim to provide maximal consistency between the recoloured image and the original [27].

Several solutions have been developed to address particular types and severities of colour blindness, such as anomalous trichromats [51, 70, 58, 71, 42].

Chapter 3

Colour in Information Visualization

Colour is now a widespread tool available to authors of information visualizations. With modern colour monitors, colour printers, and other forms of colour media, it is a rarity to find a visualization that does not use colour to express additional information (or at least for aesthetic reasons). As most forms of media limit the spatial dimensions to two, colour allows additional dimensions of data to be represented (at least one, but usually more) [68]. This increases the amount of information that can be presented and has been shown to improve performance over monochrome information visualizations [36].

Although colour reproduction quality can vary widely between devices and locations [40], authors tend to make the simplifying assumption that any colour they select will be realistically reproduced for their viewing audience. The reason for this simplifying assumption is that the alternative of anticipating all variations in colour reproduction techniques would be very difficult and time-consuming. There also exist colour profile systems which aim to provide consistent colour reproduction across devices [41].

In addition to the above-mentioned colour reproduction variations, there are many additional factors that influence how an individual user of a visualization perceives colour. These can range from genetically-caused colour vision deficiency (colour blindness), to ambient light conditions in the viewing environment, to colour perception side-effects from using prescription drugs [59]. See Chapter 2 for more details of the factors that influence colour perception. To describe the combined effects of device variation, environmental impact, and inter-user diversity, the term *atypical colour perception* is used.

In situations where information visualization users experience atypical colour perception, the author's reasonable assumptions about colour consistency are invalidated. When this occurs, difficulties in using the visualization arise. These include confusion of multiple colours that the author intended to be perceived as distinct, and a significant loss of the advantages provided by colour, as well as decreased confidence in using the visualization.

In this chapter, the uses of colour in information visualization will be outlined, as well as potential difficulties these uses can cause in atypical colour perception situations. Seven common uses of colour in information visualization will first be presented. These seven colour techniques are then analyzed to see how atypical colour perception can cause difficulties when using this techniques.

3.1 Colour use in Information Visualization

Tufte [62] describes four main uses of colour in information visualization: to *label*, to *measure*, to *imitate reality*, and to *enliven*. These four categories are reflected also in [60], and [68], with the addition of *shading to reveal shape*, and *multi-dimensional data display*, respectively. In this section, we examine several colour use techniques, and categorize them into the above classification.

3.1.1 Colour as Label

Visual preattentive processing in the human visual system allows rapid identification of colours [68]. This ability creates a special opportunity for colour as a visual element, allowing the quick identification of special, unique, or related elements. Labeling using colour allows *nominal* information to be identified quickly and efficiently. As an example, recall the colour scheme for road signs in the province of Saskatchewan. Orange signs indicate construction (speed adjustments, detours), green signs reveal general information (city distances, tourist attractions), and red signs dictate control (stop signs, yield signs). See Figure 3.1 for examples.

Some information visualization techniques that use colour as label are:

Figure 3.1: Examples of colour coding of road signs. From left to right: a stop sign, an information sign and a construction sign. From http://www.wpclipart.com, http://www.colourlovers.com, and http://www.jupiterimages.com.

- categorical encoding (e.g., separate data categories in a scatterplot)
- importance encoding - popout (e.g., find the blue boxes in a field of red boxes)
- highlighting (e.g., identify important portions of text in a document)
- brushing (e.g., distinguish data points under consideration in a scatterplot)

When using colour as label, the core task is to differentiate between colours. Categories must be discernibly different from each other and the background colour. Popout requires significant differentiability between the popout colour and all other colours. Highlighting can be implemented using subtly differentiable colours. This allows the individual using the visualization to ignore the highlighted areas if they are not important for their current work. Brushing requires colours that are substantially different from the remaining label colours.

Categorical Encoding

In categorical encoding, a unique colour is assigned to each category of data. All representations of this category in the visualization will then employ this colour as an identifying characteristic. Generally, only the colour is required in order to identify the category of an element, but additional hints (such as ordering, if possible) can help. In [21], Healy suggests a maximum of seven category colours if the luminance of the colours is held constant. With variations in luminance allowed, the number of unique categories is likely to increase. Figure 3.2 contains an example of categorical encoding, using twenty categories, which may exceed 'safe' maximum category numbers. These twenty colours are the first twenty category colours suggested by

Microsoft Excel 2004 for Mac when constructing a chart. Considerations must also be made for the background colour and any additional colours present, such as that used for label text or label backgrounds.

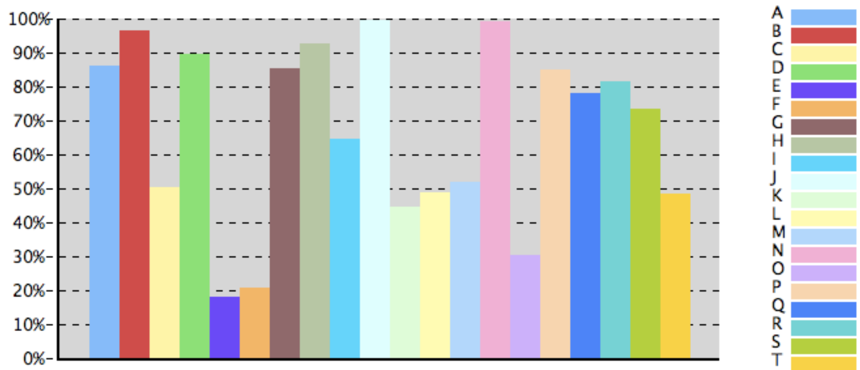

Figure 3.2: An example of categorical encoding. Each category of data is given a unique colour such that its value can be identified. Note that only a label's colour is required to identify which bar contains the appropriate information. Colours are the first twenty default colours used for chart construction in Microsoft Excel 2004 for Mac.

Importance Encoding - Popout

In popout, a unique colour is used to help identify important elements. This colour must be sufficiently different from the background colour and other colours used in the visualization. Generally, a saturated, bright, primary colour is used to replace the established element colour. The difference between the popout colour and remaining colours must be significant enough to allow preattentive identification of elements encoded using the popout colour. Due to this preattentive affect, popout allows rapid identification and location of important items. See Figure 3.3 for an example.

Highlighting

Highlighting is the use of a subtle, desaturated colour to bring attention to an element or region of a visualization. Unlike popout, highlighting does not replace the element colour in the visualization, but surrounds the element of interest. As a result, a 'soft', desaturated colour is often used to prevent the highlight from occluding (overshadowing) the highlighted item. For an example, see Figure 3.4.

```
7 9 3 8 0 2 4 8 3 9 0 5 2 2 7 3 7 9 0 2 3 9 9 7 0 3 9 8 6 5
7 6 2 7 0 3 9 9 9 1 7 2 3 6 5 5 8 1 4 7 1 3 8 4 8 0 4 6 0 3
2 6 9 4 1 3 7 8 8 3 8 1 5 3 5 4 3 6 5 9 5 4 9 1 7 5 5 4 1 8
8 3 5 2 2 6 6 7 8 4 1 7 1 8 7 8 7 7 4 4 9 1 5 5 5 8 2 9 8 2
0 7 4 8 5 8 3 0 6 2 2 5 2 2 7 1 5 2 1 1 0 1 8 7 6 0 0 9 5 6
```

Figure 3.3: An example of popout. The time required to count the number of sevens is proportional to the number of sevens. The time to count the number of threes is proportional to the total number of numbers.

```java
public static void main(String[] args)
{
    generator = new Random();
    colourFactory = new ColourFactory();
    mainFrame = new JFrame("A Multi-Bar Chart");
    mainFrame.setDefaultCloseOperation(JFrame.EXIT_ON_CLOSE);
    mainPanel = new ChartPanel();
    mainPanel.getInputMap().put(KeyStroke.getKeyStroke(KeyEvent.VK_SPACE, 0), "spacebar");
    mainPanel.getActionMap().put("spacebar", new ChangeChart());
    mainFrame.add(mainPanel);
    mainFrame.setPreferredSize(new Dimension(WINDOW_WIDTH, WINDOW_HEIGHT+MAC_FIX));
    mainFrame.pack();
    mainFrame.setVisible(true);
}
```

Figure 3.4: An example of highlighting. The variable `mainFrame` is highlighted with grey such that all occurrences can be identified. Also the line on which the cursor is located is a very subtle blue. Note that the highlighting is quite desaturated so that it does not overwhelm the rest of the code, and can therefore be easily ignored by the user.

Brushing

Brushing is a special application of highlighting to a visualization with numerous data points. Brushing is often employed as an interaction technique, in which the user of a visualization marks elements of interest so they remain easily discernible while the data is manipulated in some manner (perhaps rotated in a 3D scatterplot). As such, a brushing colour replaces the element colour, and therefore needs to be substantially different from the other colours used in the visualization. Brushing colours often are brighter than the remaining colours in the visualization, to aid in their preattentive location. Figure 3.5 contains a example of brushing in a scatterplot.

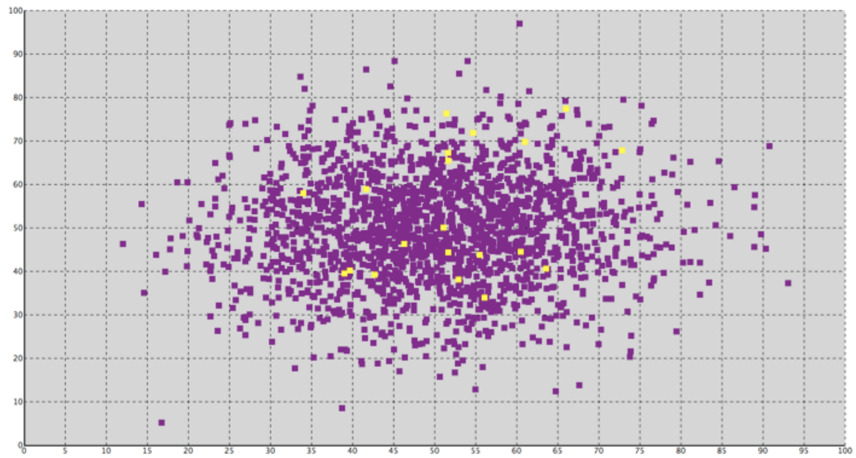

Figure 3.5: An example of brushing. In this scatterplot of two thousand data points (purple), ten have been brushed with a unique yellow colour for further consideration. Note how the brushed elements popout from the main data points.

3.1.2 Colour as Value

Colour is often used as a means to encode univariate or multivariate data. This process involves the discretization of a continuous data set. Each of the discrete ranges is associated with a given colour, and the data set is painted accordingly. For example, the depth of a body of water can be encoded using shades of blue, where darker blues indicate deeper water and lighter blues show shallower water.

Colour as measurement is used in the following information visualization tools:

- false colour representations (e.g., density as hue bands in MRI scans)
- continuums (e.g., darkness of blues to indicate ocean depth in maps)
- multi-dimensional data display (e.g., 5D data representation)

As the data set to be visualized is continuous in nature, the data set is often ordinal (i.e., orderable), and as such some ordinal properties of colours should be used to associate the data with colour. As there are three perceptual components of colour (hue, saturation, and value - see Section 2.3.3), the association between the data set and colour often occurs along one of these three components.

Continuous Data Mapping to Hue

As hue does not have any implicit ordering (Is yellow less than blue?), mapping an ordinal set of data to a nominal representation requires the invention of a hue scale. This hue scale must be learned before the visualization can be used, otherwise continual reference back to the scale is required. Maureen Stone attacks any representation in this manner [59], simply because of the artificial (and unnecessary) step of forcing the user of a visualization to learn the mapping. Ware [68] gives some credence to this approach however, by showing that frequent changes in hue can offset some interpretation difficulties. One such difficulty is simultaneous contrast, which occurs when the perception of a colour is influenced by the colours surrounding it.

As this approach maps a sequence of data points to an artificial hue sequence, it is commonly referred to as a 'false colour mapping', as each hue is perceived as a distinct colour [68]. False colour representations use an invented colour scheme to paint a set of data values. The data values are often presented as a greyscale image (wherein the greyscale intensity of a pixel represents that data point's value), and the greyscale values are mapped to the colour model to produce a false coloured image. A common hue colour scale is to use the rainbow pattern, in which the spectrum of colours is used as a sequence (red, orange, yellow, green, cyan, blue, purple). False colour representations allow specific structures to be identified in a data set (such as organs and tissues in Figure 3.6), as each variation in data produces a transition in colour, thereby outlining the structure.

Continuous Data Mapping to Saturation and Value

As both saturation and value are ordinal properties of colour (one colour can be more saturated or less bright than another colour), these attributes allow a natural mapping for ordinal data [59]. Although each property would serve this function independently, it is often difficult to hold one constant while altering the other. As such, a combined saturation-value mapping is often used, while holding hue constant (which is computationally simpler).

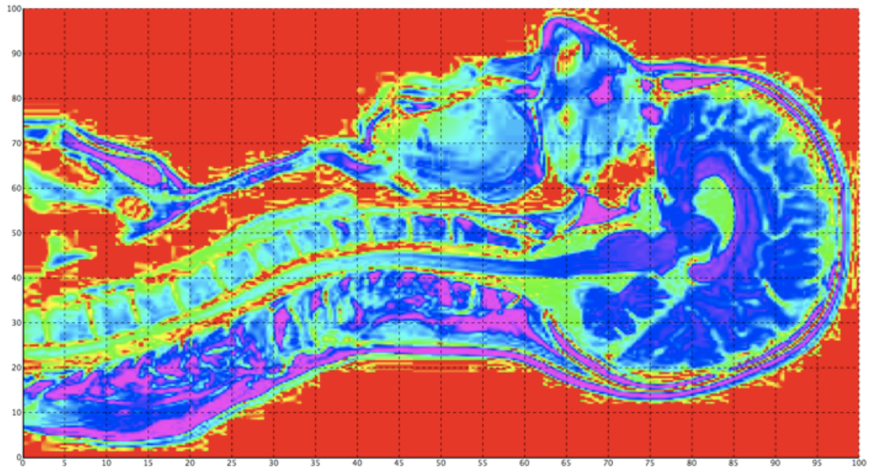

Figure 3.6: An example of false colour. This 'spectrum' order of hues (red, orange, yellow, green, cyan, blue, purple) is a common choice, although the ordering is a learned one. Original image from: `http://www.southwestimaging.co.uk`

Continuums of data allow a clean representation of trends in the data, rather than specific structure colouring (as in false colour). As the hue and/or brightness of a colour will indicate an increase or decrease in the data set values, it is simple to find local maxima and minima, areas of rapid transition, and slow gradual changes. Figure 3.7 shows a continuum of data mapped to saturation and value.

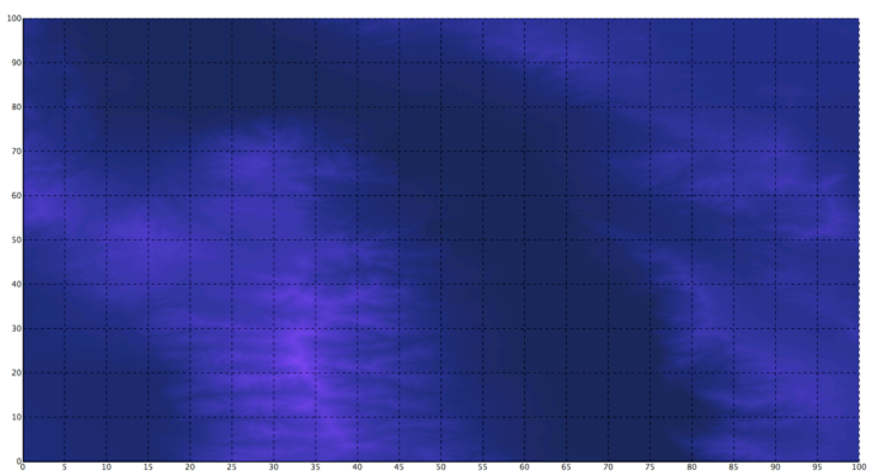

Figure 3.7: An example of continuum transition using saturation and value. As the hue (blue) is held constant, a natural continuum from dark, dull blues to light, vivid blues is present.

Multi-Dimensional Data

Multi-dimensional data can also be effectively represented using colour. This is a natural extension of continuums from a univariate space to multivariate space. Instead of using HSV models of colour, a RGB (or similar) colour model is utilized in some two-dimensional display. Three dimensions of the data can be mapped to each of the R, G, and B channels of a colour. Two additional dimensions can be represented by the (X,Y) spatial coordinates on the display. This allows five dimensions of data to be represented [69]. Although these may be an effective means of encoding multi-dimensional data [52], they may be quite difficult to understand [65].

As a clarifying example, imagine a visualization presented as a map of a municipality. Each square kilometer of the municipality is encoded using an appropriately-scaled square on the map. Each of these squares encodes three pieces of information: average income of all residents of the square, total population of the square, and number of trees in the square. For the entire map, there will be a maximum and minimum for each of these three pieces of information. If average income is encoded using the RGB red channel, then the minimum average income would be red=0 (or close to 0) and the maximum average income would be red=255. A similar approach can be used for total population (green), and number of trees (blue). These three pieces of information could be combined into a single RGB colour and this could be used to colour the representative square on the map. For a given location in the municipality (two dimensions), three more pieces of information are available by interpreting the colour for its representative square (namely average income, population, and number of trees). This use of colour allows encoding of five dimensions of data on a two-dimensional map.

3.1.3 Colour as Imitation of Reality and as Decoration

In cartography or scientific visualization, real-world data is presented. Often this real-world data has a natural colouring scheme that can be carried over to its representation. In cartography, water is often shades of blue, vegetation shades of green, mountains shades of grey (with white peaks), etc. This approach encapsulates the visualization technique of using colour to imitate reality. Beautification of a visualization often involves colouring it in terms of some natural or aesthetic model to increase its appeal. These concepts also somewhat encapsulate Stone's shading extension to Tufte's list. These techniques will not be discussed further in this paper, but a convincing example of these concepts is shown in Figure 3.8.

Figure 3.8: An example of colour used for decoration and to imitate reality. Mountain peaks are white from snow, trees are green, rock outcroppings are blueish. From: `http://www.destination360.com/north-america/us/washington/crystal-mountain.php`

3.2 Effect of Atypical Colour Perception on Colour Use in Information Visualization

In this section, the effects that atypical colour perception (Section 2.4) have on information visualization colour use techniques are examined. All simulations of colour blindness are provided by the VisCheck system (`http://www.vischeck.com/`).

3.2.1 Colour as Label

When examining the uses of colour as a label (Section 3.1.1), four techniques were identified:

- categorical encoding
- popout
- highlighting
- brushing

In this section, each of these techniques is examined in order to illustrate specific problems arising from atypical colour perception.

Categorical Encoding

All types of atypical colour perception will have some level of difficulty with categorical encoding. The gamut of colours perceptible by these individuals is reduced from typical colour perception. Colours for a visualization are chosen because of their distinctiveness to non-colour blind designers. These same colours may not be distinct in atypical colour perception situations, thereby leading to confusion. As an example, Figure 3.9 contains the bar chart from Figure 3.2 simulating the colours perceived by an tritanopic individual.

As each variety of atypical colour perception results in different colours becoming indistinguishable, it is very difficult to produce a 'safe' categorical colour encoding. Colours that differ only in their greyscale value may work, except that there are severe simultaneous contrast and Mach banding issues with greyscale colours, limiting the number of effective greyscale colours to four or five [68].

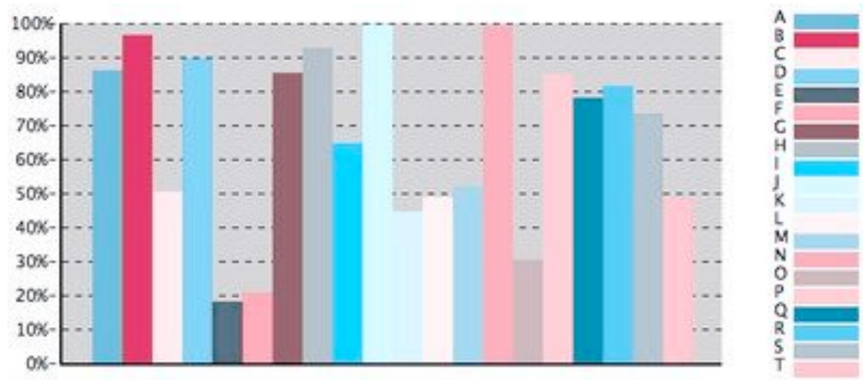

Figure 3.9: Tritanopic simulation of a bar chart categorical encoding. Note that the categories are no longer uniquely identifiable by colour alone.

Popout

The effective use of hue, saturation and/or value contrast allows popout to be an effective visualization technique. This contrast must exist between certain data points and the remaining colours utilized in the visualization in order for popout to work effectively. As can be seen in Figure 3.3, popout is an effective means of discerning important information from a large data set.

When there is insufficient contrast between the highlighted and remaining data, popout fails to help. This difficulty rarely occurs when both the author and reader have typical colour perception. In the case where the reader has atypical colour perception, however, trouble can arise. Since 'primary' colours are often used in popout situations (black and white with a bright, richly saturated hue such as red, blue, green, or yellow), there is little chance of two colours being confused, but the protanopic darkening of red colours described in Section 2.5.3, can cause difficulties with black and red. Red is often used as a popout colour as it generally grabs the attention of the viewer very effectively. [1] Unfortunately, protanopic and protanomalous individuals have a darkened response to red, causing it to be indistinguishable from dark grey or black. This causes the popout visualization for typical colour perception in Figure 3.10 to be nullified, as shown in Figure 3.11.

[1] Red is used in emergency exit handles, fire trucks, stop signs, exit signs, yield signs, stop lights, fire extinguishers, etc.

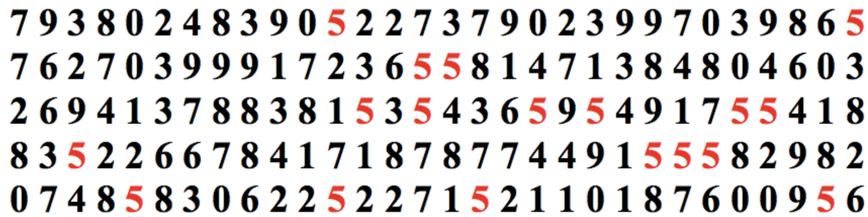

Figure 3.10: Popout example using a highlight colour of red on the fives. The fives should clearly be distinguishable from the remaining values, and countable in time proportional to their number.

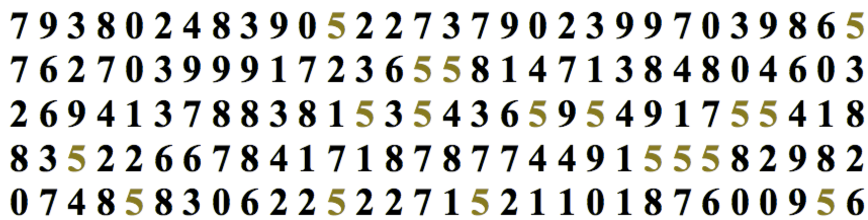

Figure 3.11: Protanopic simulation of Figure 3.10. Note how the fives are now greatly diminished in the intensity of the red, almost appearing a dark grey.

Highlighting

Highlighting requires the use of bright, desaturated colours as it is often used as an accent in situations where there is black text on a white background. If the highlighting colour it too dark, it may obscure the underlying text. A consequence of using light, desaturated colours is that they are similar in nature to white, and can be confused with white in atypical colour perception situations. Examples of these highlighting colours can be obtained by slightly reducing one of the RGB channels of pure white, as seen in Figure 3.12.

If the highlighting colour is not easily discernible from the background colour, the advantage of the highlight is obviously lost. Unfortunately, highlighting is not often used with *redundant encoding*, in which an additional visualization indicates the relevance of the highlighted area. When no additional hints are provided, an individual with atypical colour perception must search (often linearly) to find the highlighted information. An example of this is the highlight used to indicate the cursor's current vertical position in Figure 3.4. If the highlight is not visible, the

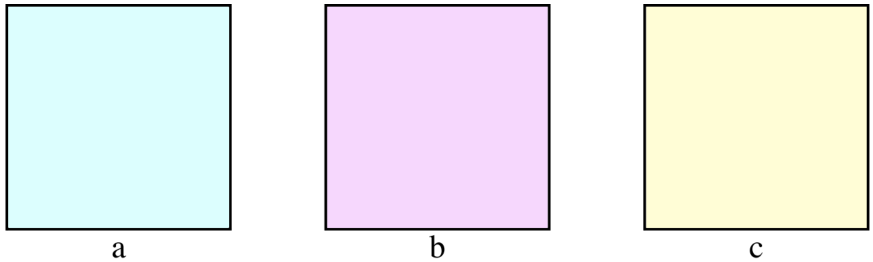

Figure 3.12: Examples of common highlighting colours. Colour 'a' was obtained by reducing the red channel to obtain (200,255,255), 'b' was obtained by reducing the green channel to obtain (255,200,255), and 'c' was obtained by reducing the blue channel to obtain (255,255,200).

reader must search from the top of the screen until the cursor is found. [2]

Highlighting colours can often be confused with the background white colour in atypical colour perception, as the difference between white and the highlight colour lies along the stimulation of a single cone type. For example, magenta (Figure 3.12.b) is produced by slightly reducing the amount of green. Green wavelengths largely influence the medium wavelength sensitive cone. If this cone behaves differently (as in deuteranopic types of colour blindness), this subtle reduction from white may not be perceptible. This problem extends to cyan and yellow for protanopic and tritanopic individuals as well.

Brushing

Many of the difficulties of popout visualization techniques also apply to brushing techniques. One complication is the potential for more colours to be present in the visualization. As different groups of data may be categorically encoded, choosing a distinct and highly visible brushing colour may be very difficult. The scatterplot example given in Figure 3.5 also uses a unique colour for the chart background (grey), in addition to the label text (black), and label background (white). This imposes more constraints on brushing colour selection.

The simple scatterplot brushing example shown in Figure 3.5 exemplifies some

[2] Luckily, the blinking nature of the cursor provides a preattentive tool (and redundant encoding) for locating it.

difficulties for atypical colour perception. The highlight colour needs to be distinct from the other data points as well as the background colour. If the non-brushed data points are coloured to represent various categories of data, it is likely that there will be some colour confusion between either the categories themselves, the category colours and the brushing colour, or the brushing colour and the background. Once again (like popout), the user is no longer able to rely on speedy preattentive processing to identify the brushed datapoints, but must exhaustively search all areas of the scatterplot to identify the brushed elements.

3.2.2 Colour as Value

When examining the uses of colour as a measurement (Section 3.1.2), we identified three key techniques:

- false colour representations
- continuums
- multi-dimensional data display

In this section, these techniques are examined in light of atypical colour perception to illuminate potential difficulties.

False Colour

As false colour techniques for mapping continuous data often pose problems in typical colour perception situations, it is easy to understand that this approach also presents significant difficulties for individuals who experience atypical colour perception. As hue is not an ordinal property, a hue ordering must be created to facilitate the mapping of the continuous data to the false colours. This mapping must also be learned by users of the visualization, which poses many problems in atypical colour perception situations.

A common false colour mapping is the use of a spectrum of colour at a relatively high brightness and saturation (e.g., Figure 3.6). Dichromatic individuals will have a particularly difficult time with this mapping, as some of the colours in the mapping will differ only in the amount they stimulate a given cone. As this may be the missing cone type, these variations will be absent. This is demonstrated for a deuteranopic viewer of the false colour image referred to above in Figure 3.13.

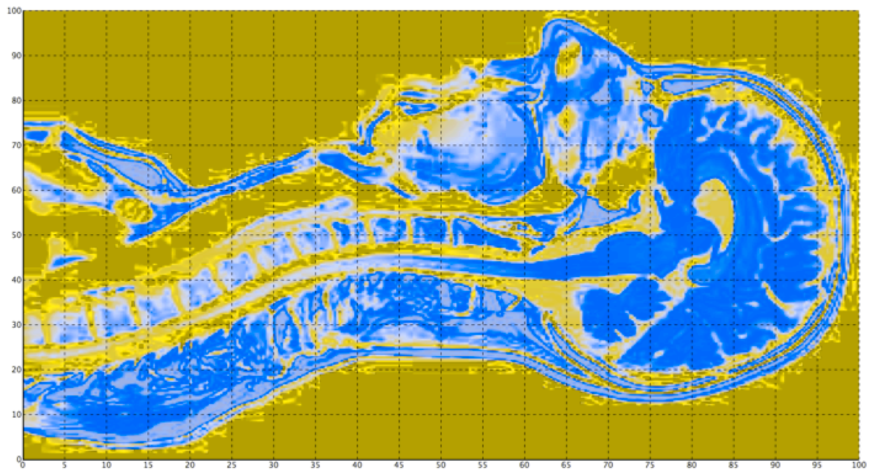

Figure 3.13: Simulation of a deuteranopic view of a false colour mapping for MRI density (Figure 3.6). Note the loss of colour uniqueness.

Likewise, if a colour is used that differs from pure white in the manner described in Section 3.2.1, these colours will be confused with white (which is a natural endpoint of a mapping). As a result, areas are identified as extremes in the data when they should not be. As an example, the cyan (green-blue) colour in Figure 3.6 appears as white to protanopes, thereby mislabeling areas of medium density as high density.

Continuums

As outlined in Section 3.1.2, discrete representation of continuous data often utilizes nominal properties of colour such as saturation or lightness (or both). These schemes tend to work well in both typical and atypical colour perception, as the variations in brightness and/or saturation can even be detectable by individuals with monochromatic colour vision.[3]

[3] Although the notion of saturation should be foreign to monochromats, the information encoded as saturation differences should be perceived as changes in lightness (as long as saturation and lightness are not encoding two separate dimensions of data.

Unfortunately, hue is also used to represent continuous data, even though it lacks any simple representation of order. A common scheme is to use a transition from one colour to another colour (e.g., like blue for cold temperatures and red for hot temperatures, with colours 'between' them representing temperatures between cold and hot). This poses many difficulties in atypical colour perception situations. As an example, Figure 3.14 contains a commonly used transition from red to green. As red and green are quite distinct colours for those with typical colour perception, this works well for these individuals. The most common forms of colour blindness (protanopic and deuteranopic problems) cannot use this colour scheme effectively. Protanopic and deuteranopic viewer simulations of the same colour scheme are in Figures 3.15 and 3.16, respectively. It is easy to see that the distinct transition from red to green is missing in both. This will cause severe difficulties when these users are trying to 'read' information from a visualization that utilizes this colour scheme. This problem carries over to tritanomalous or tritanopic viewers of yellow to blue transitions.

Figure 3.14: A continuous colour scheme that works very poorly for those with red-green colour blindness.

Figure 3.15: The same continuous colour scheme as it appears to a protanopic viewer.

Figure 3.16: The same continuous colour scheme as it appears to a deuteranopic viewer.

Multi-Dimensional Data

The three-dimensional RGB colour space used in multi-dimensional data visualization is an analogue for the three-dimensional XYZ cone stimulus model presented in Figure 2.8. As such, any compression of the colour perception cube as a result of atypical colour perception will reduce the volume of the comparable RGB colour cube. In general, this will reduce the effectiveness of this visualization tool, as the range of values (most likely in one of the three channels of RGB) will be significantly reduced. This simply renders the visualization of that specific dimension of data to be nearly useless, thereby reducing the effectiveness of the visualization and increasing the chance for error.

Chapter 4

A Model of Context-Specific Colour Differentiation

In this chapter, the high-level ideas of the model presented in this thesis are described. This model is based on a working definition of colour differentiation.

4.1 What is colour differentiation?

Colour differentiation is the ability of an individual to identify that two colours are different. Some colours may be very different from each other (such as yellow and blue), and other colours may not be very different (yellowish-green and greenish-yellow). This suggests that for any given colour, there will be colours that are perceptibly 'close' to it (not very different), and colours that will be perceptibly 'far' away (very different).

In this section, this definition of colour differentiation is developed further. This requires a description of RGB colour space, and the notion of *channels* within RGB colour space.

4.1.1 RGB colour space

In digital environments, the properties of colour are expressed in terms of a given colour model. Most frequently, this model is a standardized model called RGB, which produces colours using a blend of three primary light sources, one red, one green, and one blue (hence RGB). The electro-magnetic radiation (EMR) frequency of each of these primary light sources are selected in consideration of the three colour photoreceptors of the retina, such that the red primary stimulates the long-wavelength cone, the green primary stimulates the medium-wavelength cone, and the blue primary stimulates the short-wavelength cone. For colour production in digital environments (generally through some manner of monitor), the three primary light sources are combined into a single pixel, such that the individual primary light sources are not distinguishable from one another. This combination allows each pixel to selectively stimulate the three types of cones of the retina, and thereby induce the perception of colour.

RGB is an *additive* model of colour, in which each individual component of colour production adds to the overall colour. In the case of RGB, increasing the 'R' primary source light increases the amount of red in the overall colour, increasing the 'G' primary source light increases the amount of green in the overall colour, and increasing the 'B' primary source light increases the amount of blue in the overall colour. When the three primary light sources are producing no output, the colour perceived is black (non-stimulation of cones). When all three primary light sources are producing maximum output, the colour perceived is white (maximum uniform stimulation of cones). Varying which primary lights are producing maximum output and which are not produces the eight principle colours of digital systems (three primary colours - red, green, blue; three secondary colours - yellow, purple, cyan; black and white), as is shown in Table 4.1.

As is suggested by the terms 'no output' and 'maximum output' used above, each primary source light is able to produce a wide range of output values between these two extremes. Most commonly, the continuous range of output for each primary source light is discretized into 256 steps. These values (ranging from 0-255), are encoded using a single byte (eight bits) of memory. As a result of this discretization, every colour produceable in a digital environment can be encoded as a triple of values, with each value in the range [0-255].

As there are three primary colours, RGB colour can be represented as a three-dimensional volume, where the value of each of the primary sources is encoded as one of the dimensions. In this thesis, this volume is known as the RGB colour space, and

Red State	Green State	Blue State	Resulting Colour
Max output	No output	No output	Red
No output	Max output	No output	Green
No output	No output	Max output	Blue
Max output	Max output	No output	Yellow
Max output	No output	Max output	Purple
No output	Max output	Max output	Cyan
No output	No output	No output	Black
Max output	Max output	Max output	White

Table 4.1: Eight possible primary colours produced in RGB colour system.

is illustrated in Figure 4.1. Each of the individual primary source lights are known as *channels*, such that there is a red channel, a green channel, and a blue channel, each having possible values in the range [0-255].

4.1.2 Colour differentiation

As discussed above, a specific colour can be encoded as a triple of values in the range [0-255]. This triple specifies a particular location in the RGB colour space shown in Figure 4.1. Two distinct colours will have different locations in the RGB colour space. This notion of different locations for different colours allows a formal specification of colour differentiation.

Let C_1 be a colour with RGB values $R_1G_1B_1$. Let C_2 be a different colour, with RGB values $R_2G_2B_2$. The relationship between these two colours can be described by a triple of values $N_RN_GN_B$ defined as follows:

$$|R_1 - R_2| = N_R$$

$$|G_1 - G_2| = N_G$$

$$|B_1 - B_2| = N_B$$

Figure 4.1: RGB colour cube, with red, purple, white, and yellow in the front, and black (hidden), blue, cyan, and green in the back. From: `http://www.cs.ru.nl/~ths/rt2/col/h2/2fundENG.html`

where $N_R + N_G + N_B = D_M$, the Manhattan distance between the two colours in RGB colour space. Although the Manhattan distance between two colours is an appealing estimate for colour differentiation, a simpler description of colour differentiation is put forward here.

Suppose colour C_1 with RGB values $R_1G_1B_1$ and colour C_2 with RGB values $R_2G_2B_2$ are perceptibly different (i.e., they look like different colours). It is hypothesized that this perceptible difference results from at least one of the following conditions being true:

1. $N_R > L_R$
2. $N_G > L_G$
3. $N_B > L_B$

where L_R represents some threshold limit for the red channel, L_G for the green channel, and L_B for the blue channel. When a value for N exceeds its respective limit L, it is assumed that the two colours involved are differentiable. What this essentially means is that if two colours have a sufficient difference between any of their respective channels, then they will be perceived as different. On the other hand, if two colours do not differ sufficiently between any of their respective channels, then they will not be perceived as different. The values L_R, L_G, and L_B encapsulate this notion of 'sufficiently different'.

There is a subtlety hidden by the use of absolute value in the formalism given above. If the absolute value operation is eliminated, the following results:

$$R_1 - R_2 = N_R$$

$$G_1 - G_2 = N_G$$

$$B_1 - B_2 = N_B$$

in which N_R, N_G, and N_B can now be either positive or negative (represented with an additional $+$ or $-$ in the subscript). Thus positive and negative values for L_R, L_G, and L_B would be required (again represented with an additional $+$ or $-$ in the subscript), in which differentiation is achieved when at least one of the following conditions is true:

1. $(R_1 - R_2 = N_{+R} > L_{+R}) \ \ OR \ \ (R_1 - R_2 = N_{-R} < L_{-R})$

2. $(G_1 - G_2 = N_{+G} > L_{+G}) \ \ OR \ \ (G_1 - G_2 = N_{-G} < L_{-G})$

3. $(B_1 - B_2 = N_{+B} > L_{+B}) \ \ OR \ \ (B_1 - B_2 = N_{-B} < L_{-B})$

Now determining the values for $L_{\pm R}$, $L_{\pm G}$, and $L_{\pm B}$ can be considered. These values are collectively referred to as the *limits* for a particular colour. Given a colour, C_X, with values $R_X G_X B_X$, the limits are defined as a sextet $L_{+R}L_{-R}L_{+G}L_{-G}L_{+B}L_{-B}$. It is assumed that the values of R_X, G_X, and B_X influence the values for the limits, and this is explored further in Chapter 6. Additionally, the environment in which the colours are perceived will also influence the values of the limits. In this thesis, an environment is described as a *context* and is the union of all factors that influence colour perception, and is explored next.

4.2 What is Context-Specific Colour Differentiation?

As discussed in Chapter 2, two major categories of factors that influence colour perception exist: external and internal. External factors are the factors that influence colour perception in the external environment. These include the qualities of the source light (brightness and frequency distribution), and the presence of filtering materials, and any reflection from external surfaces. Internal factors are the factors that influence colour perception that are internal to the perceiver of the colour. These include the presence of eye disease (e.g., cataracts), colour blindness, deterioration of the retina (e.g., from long-term diabetes), and neurological disorders and fatigue.

In this thesis, the combined effect of every external and internal factor is called a *context*. Constructing a model of colour perception that is sensitive to the context of the perception must consider every external and internal factor that influences colour perception. To do this, two approaches can be taken. The first is to identify every single factor of colour perception, and control for each one in the model. Although this would lead to a thorough model, it would be a difficult approach for many reasons. First, it is impossible to know that every factor has been identified. The number of factors is unknown and possibly unknowable, and identification of them would rely on much more advanced theories of colour perception than are currently available. Second, the interaction between factors (such as light sources and indirect surface reflectance) would dramatically complicate the modeling process, as controlling for one factor could possibly alter the control technique for other factors.

The second approach to considering every external and internal factor that influences colour perception is to use a judgement task to identify how an individual perceives colour. If the judgement task is performed in the very context that is being considered, then every internal and external factor will be automatically included in the model, as they are at work influencing the judgement task in the same manner as they influence real-world tasks in the given environment. This is the approach taken in this thesis.

In summary, context-specific colour perception is the perception of colour in a given environment. An environment contains numerous external and internal factors that influence colour perception. A *context* encapsulates all of these factors for a given environment. In this research, a context includes all these factors, but the source of colour is a monitor or projector which is driven by a digital computing system.

4.3 A Model of Context-Specific Colour Differentiation

Every location in the RGB colour space is a spatial representation of a unique colour. Two colours are considered differentiable if they differ by at least a certain amount along at least one channel. This difference is called a limit. Colour perception is influenced by many factors, both internal and external. As a side-effect of this influence on colour perception, these factors also influence colour differentiation. As a result, the limit values for a particular colour may not be the same as the limit values for another colour.

The model of context-specific colour differentiation is expressed in terms of the RGB colour space. For each colour $C_X = R_X G_X B_X$ in the RGB colour space, there is a sextet of values $L_{+R}L_{-R}L_{+G}L_{-G}L_{+B}L_{-B}$ that define the limits above and below R_X, G_X, and B_X along each channel. Any colour outside of the range defined by these limits is perceived as different by a particular individual in a particular context.

4.4 Are there Relationships Between Limits?

Each colour in RGB colour space has associated with it a set of six values that define the upper and lower differentiability limits along each channel. In Section 4.2, it was described how the context of colour perception influences how the colour is perceived, as well as differentiation between two colours. As a context includes the environment, the colour producing technology, and the individual perceiving the colour, it stands to reason that each context has a unique set of limits for the colours of RGB colour space. In this section, the relationships between the limits in a specific context will be examined to identify any relationships between the limits. Identification of these relationships will allow a simplification of the model of context-specific colour differentiation.

Early evidence suggested that there is indeed some relationship between the limits for a particular individual in a particular environment. This early evidence was produced in pilot and small-scale studies I performed. The purpose of these studies was to measure colour differentiability in constrained situations. These situations often involved holding one or two of the channels constant (often at zero), and measuring differentiability within the remaining channel. One sample of these early experiments is shown in Figure 4.2. In this figure, a clear linear relationship ($R^2 = 0.98$) exists in which the upper differentiability limit (L_{+G}) increases linearly with the green channel start value. The red and blue channels were held constant at zero for this study, which only involved one participant.

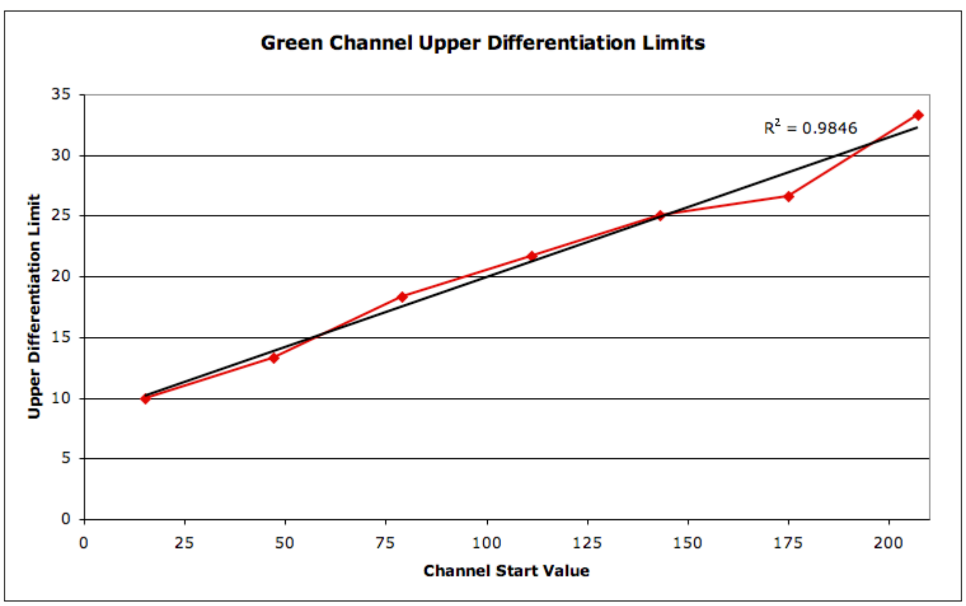

Figure 4.2: Green channel differentiability. Clear linear relationship present with R^2 of 0.98 between the green channel start value and differentiable limit. These results are based on a single participant.

Figure 4.3 illustrates another early finding. Using greyscale colour only (i.e.,

$R = G = B$), differentiability limits were again measured. The findings revealed a clear linear relationship again ($R^2 = 0.99$), between greyscale start value and the differentiability limit. This study was again a small-scale study of just one participant.

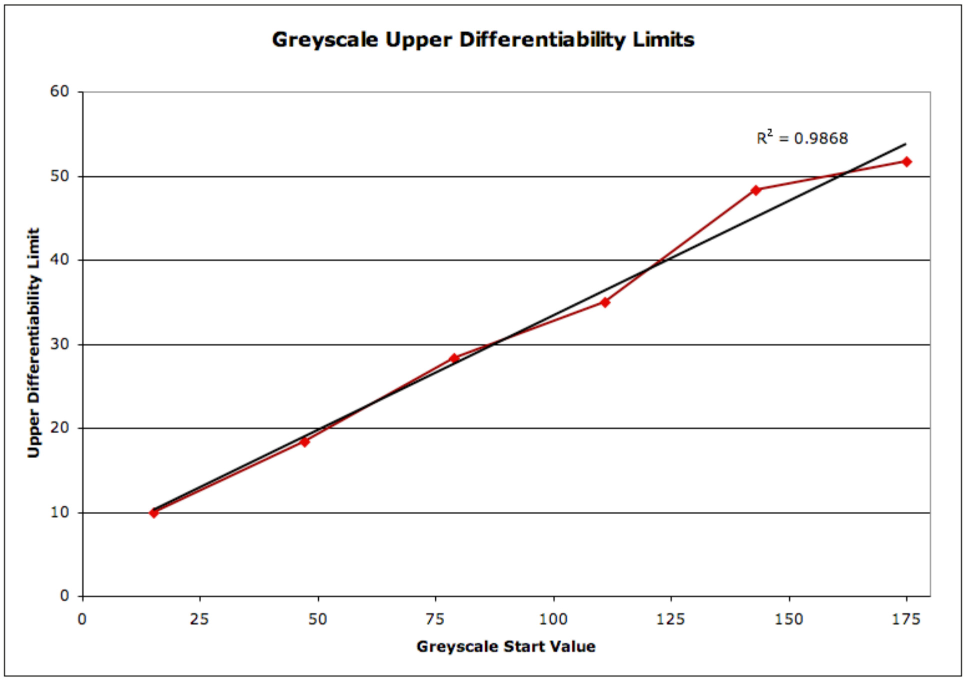

Figure 4.3: Greyscale differentiability. Clear linear relationship present with R^2 of 0.99 between the greyscale start value and differentiable limit. These results are based on a single participant.

As a third and final result of these early studies, Figure 4.4 presents results from a single participant study performed in which the testing channel (in this case, red), was held constant, but the remaining channels fluctuated. The remaining channels (green and blue) were always identical to each other, but varied uniformly from 0 to 255 in value. The value of the green and blue channels is plotted against the upper differentiability limit (L_{+R}) for the red channel, which always had a starting value of zero. A clear linear relationship is present ($R^2 = 0.97$) between the green and blue channel value and the differentiability of the red channel.

These small-scale preliminary studies revealed that there may be linear relationships present between the values of L. To illustrate this idea further, imagine shooting rays through the RGB colour cube. Each ray will intersect with a line of

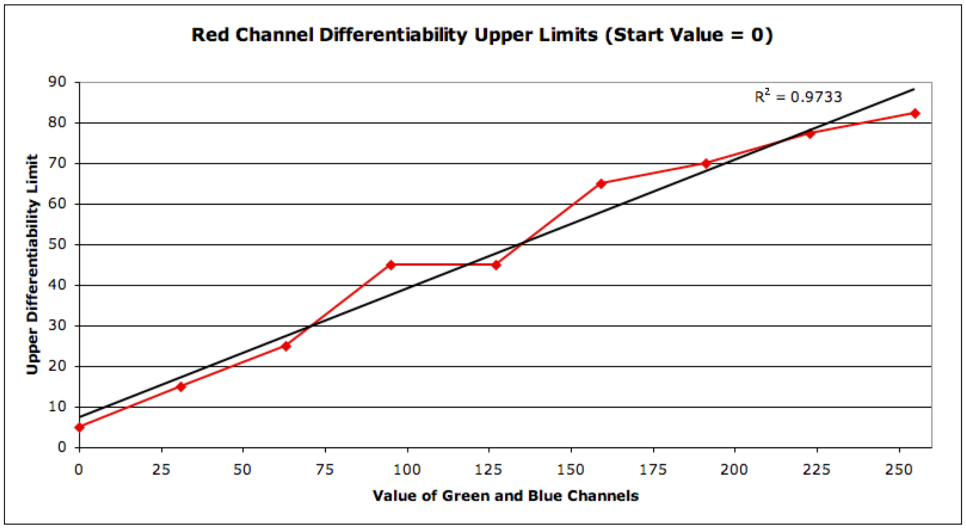

Figure 4.4: Red channel differentiability when green and blue channels vary between 0 and 255. Clear linear relationship present with R^2 of 0.97 between the green and blue channel values and and the red channel differentiable limit. Red channel start value remained at zero for entire study. These results are based on a single participant.

RGB colour values in which at least one of the channels is changing. These early results suggest that the limits for each colour along this line vary in a linear fashion to the RGB colour values along the line. This hypothesis is the subject of Chapter 6, in which a larger-scale study is performed to evaluate the potential of linear functions to explain the relationships between limits.

4.5 Eliminating Lower Limit Measurements

Establishing both lower limits and upper limits would require many measurements. If one of these could be eliminated, the number of measurements could be reduced by half. This is indeed possible, and the elimination of lower limit measurements is given here.

Presume that the upper limits for every colour in the RGB colour cube are known, and the lower red channel limit is wanted for a particular colour. To determine the lower limit for this `colour`, the following algorithm can be used. This algorithm assumes that there exists a `Colour` object that encodes the information for a RGB colour, and the function `findUpperLimit` which accepts a colour and a string encoding the channel for which the upper limit is desired, and returns the upper limit for the channel and colour provided.

```
int origRed ← colour.getRed()
int origGreen ← colour.getGreen()
int origBlue ← colour.getBlue()
int red ← origRed
int redUpperLimit
do
    red ← red - 1
    Colour tempColour ← new Colour(red, origGreen, origBlue)
    redUpperLimit ← findUpperLimit(tempColour, "red")
while (red + redUpperLimit > origRed)
int redLowerLimit ← redUpperLimit
```

This algorithm proceeds down the red channel to find new colours. The upper limit for the red channel for each of these new colours is found. When this upper limit added to the new colour red channel value is less than or equal to the original red channel value, then the lower limit has been found. This is based on the fact that for a particular channel, the lower limit for a given colour is also the upper limit for a colour that has a lower value for the particular channel. This can be likewise implemented for finding green and blue channel lower limits.

Chapter 5

Implementation of the Colour Differentiation Model

In this chapter, the implementation of the colour differentiation model presented in Chapter 4 will be discussed.

5.1 Assumption of Linearity

As mentioned in Section 4.4, the presence of a simple mathematical relationship to describe colour differentiation limits over the entire RGB colour space allows a simpler model. For the purposes of this implementation, it is assumed that the linear relationship presented in Section 4.4 holds.

The assumption of linearity allows a dramatic simplification of the general model presented in Chapter 4. In that model, every colour in the RGB colour space had associated with it a sextet of values, $L_{+R}L_{-R}L_{+G}L_{-G}L_{+B}L_{-B}$. These values are the colour differentiation limits for the colour, one for above and one for below each channel. To construct a model of colour differentiation for a given context, the limits for every colour in the RGB colour cube would need to be measured. As there are six limits per colour, and 16,777,216 possible colours[1], this approach would require $6 \times 16,777,216 = 100,663,296$ measurements that would need to be made. If each measurement took one second, the entire process would require almost 3.2 years! It is easy to see that this approach is impractical. Using the linearity assumption, the number of measurements can be reduced from 100,663,296 to 24. Achieving this

[1] 3 channels at 256 (2^8) values per channel $= 2^8 \times 2^8 \times 2^8 = 2^{24}$

reduction is explored now.

Tracing an arbitrary ray through the RGB colour cube results in two sequences. The first sequence is the colours the ray intersects. The second sequence contains the sextet of limits for each colour. As a ray is straight, at least one of the channels in the sequence of colours is changing in a linear manner (for each step along the line, at least one channel increases or decreases by a set amount). The linear assumption states that the values of the second sequence (limits) varies proportionally to the values of the first sequence. This means that as the ray is traversed, the limits along the ray change in a linear fashion (in a manner directly proportional to the change in the colour channels along the ray). To illustrate this, Figure 4.2 is presented here again as Figure 5.1. In this figure, the ray passes between two corners. It begins in the corner (0,0,0) - black, and progresses to the corner (0,255,0) - green. At select points along this path[2], the upper limit for the green channel has been measured and is reported. It can be seen that the limit is directly proportional to the green channel start value. In this case, this means that as the channel start value increases, a larger difference is required before the two colours are differentiable.

5.2 One-Dimensional Case

Any linear function can be fully described using two points ($P_{x1,y1}$, $P_{x2,y2}$) along the line. Using these two points, the slope (m) and intercept (b) of the linear function can be calculated:

$$slope = \Delta y/\Delta x = (y2 - y1)/(x2 - x1) = m$$

$$intercept = y - mx = y1 - m(x1) = y2 - m(x2) = b$$

With the definition of the linear function, any value along the linear function can now be calculated using the processes of linear interpolation and linear extrapolation. These processes are simplified by the knowledge of m and b, such that to determine

[2] red channel = 0, blue channel = 0, green channel $\in \{15, 47, 79, 111, 143, 175, 207\}$

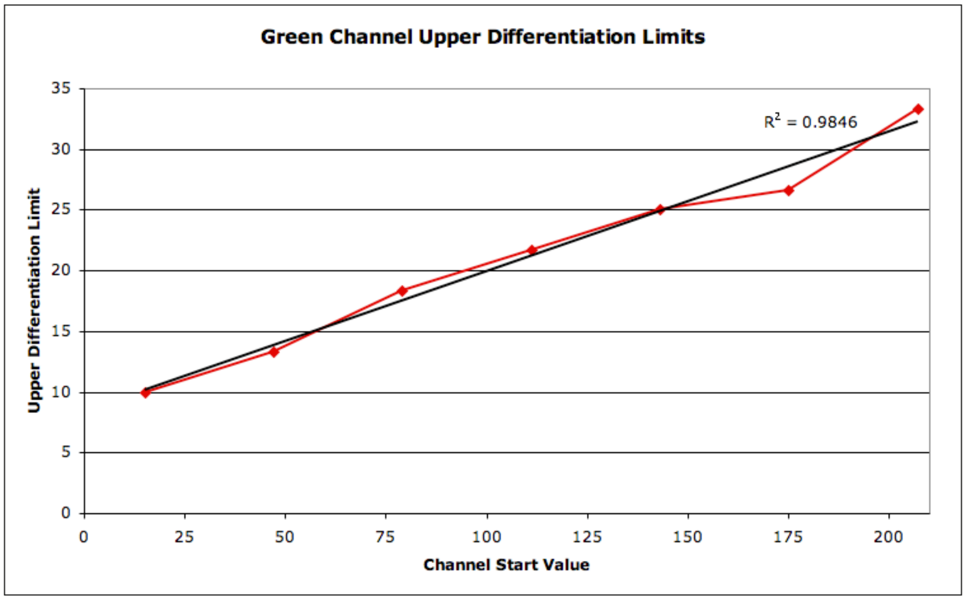

Figure 5.1: Green channel upper differentiability limit along the ray from the black corner of the RGB colour cube (0,0,0) to the green corner (0,255,0).

any dependent variable value (y) for a given dependent variable value (x), we simply solve the formula $y = mx + b$.

Applying this to the model of colour differentiation by using the linear assumption allows all of the limits along a line in the cube to be determined by making two measurements of limits along the line. Imagine a simplified differentiation model that only examines discernment along one channel (red), and only along one outside edge of the RGB colour cube (say from black (0,0,0) to red (255,0,0)). If two measurements of red differentiability (two limits) can be made along this line, the differentiability limits along the entire line can then be calculated using the formula given above.

The channel that is changing along this edge is the red channel, so this alone serves as the independent variable for the function. The differentiability limit along this edge serves as the dependent variable. Now imagine that the two measurements are made at the ends of the edge (i.e., at (0,0,0) and (255,0,0)), and the upper limits are 15 and 89, respectively. Then:

$$slope = (89 - 15)/(255 - 0) = 74/255 = 0.29 = m$$

$$intercept = y - mx = y1 - m(x1) = 15 - (0.29 \times 0) = 15 = b$$

Now any limit along this edge can be calculated from the values m and b determined above by the following formula:

$$limit = (0.29 \times red\_channel\_value) + 15$$

5.3 Two-Dimensional Case

The approach described in Section 5.2 can be extended to the two-dimensional arena by applying the linearity assumption. Imagine a less-simplified differentiation model that still only examines discernment along one channel (again, red), but along an outer face of the RGB colour cube (say the face defined by the four corners: black (0,0,0), red (255,0,0), green (0,255,0), and yellow (255,255,0)). If measurements of red differentiability are made at these four corners, then two linear functions can be generated, and used to calculate the red differentiability limit at any given point on this face, defined as $P_{Red,Green}$.

Even though four limit measurements can generate four potential functions (Black-Red, Black-Green, Red-Yellow, and Green-Yellow), only those that vary along the red channel axis are useful (Black-Red and Green-Yellow). This is directly linked to the fact that only red limits are under consideration in this limited model. When green or blue limits are under consideration, then the functions that vary along their respective channel axes are useful.

Each edge function is generated, and then the red coordinate of the point under consideration is fed into each function to generate two new limits. These two limits can then be used to generate another function (in the same process as the first two functions). This function is then fed the green coordinate of the point under consideration to calculate a single red differentiation limit.

The process is:

1. Measure red differentiability limits at each corner: Black = L_B, Red = L_R, Green = L_G, and Yellow = L_Y.

2. Determine Black-Red function: $F_{B-R} : y = (L_R - L_B)/255 \times x + L_B$

3. Determine Green-Yellow function: $F_{G-Y} : y = (L_Y - L_G)/255 \times x + L_G$

4. Feed the red coordinate of the point of interest ($P_{Red,Green}$) into each of the above functions to calculate two new limits: $L_{B-R(Red)}$ and $L_{G-Y(Red)}$

5. Using these two new limits, determine another linear function: $F_{B-R+G-Y}$: $y = (L_{G-Y(Red)} - L_{B-R(Red)})/255 \times x + L_{B-R(Red)}$

6. Feed the green coordinate of the point of interest ($P_{Red,Green}$) into this new function to calculate a single limit. This is the red differentiation limit for the colour specified by the point $P_{Red,Green}$ on this face of the RGB colour cube.

5.4 Full Three-Dimensional Case

To extend the 2D process to the three-dimensional arena, differentiability limits are measured at all eight corners of the RGB colour space. These eight limits are then used to define four functions along the outside edges of the cube. The edges of importance are those that vary in the channel for which the limit is desired. The point of interest (for which the limit is desired) is specified using a red, a green, and a blue coordinate. These four functions are then fed the coordinate for the channel for which the limit is desired. This gives back four limits, which can be used in the same manner as the two-dimensional description given above.

When the three-dimensional process gets to the point of using the four points to generate two new functions, as per the two-dimensional description given above, a choice needs to be made. This is illustrated by example. If the desired limit is along the blue channel, then the blue coordinate of the point of interest is used to calculate the four points. After this, two choices are available: 1) the green coordinate can be used to generate the single linear function which is fed the red coordinate to generate the desired limit, 2) the red coordinate can be used to generate the single linear function which is fed the green coordinate to generate the desired limit. It stands to reason that both approaches are equivalent, so one or the other can be selected arbitrarily. To test that this equivalence is true, both approaches were computed in the final software implementation and the results of each were compared for equality inside the following three assertion statements:

```
assert Math.abs(upperRedLimitGB - upperRedLimitBG) < 0.001
assert Math.abs(upperGreenLimitRB - upperGreenLimitBR) < 0.001
assert Math.abs(upperBlueLimitRG - upperBlueLimitGR) < 0.001
```

In the extensive number of runs of the system (thousands of predictions using this linear function approach), no assertion has ever failed.

To extend this three-dimensional model to calculating differentiability limits along all channels, the process simply needs to be repeated three times, one for each desired limit. Each iteration uses a different set of RGB colour cube corner differentiability limits. Each of these sets contain eight corner measurements, and there are three channels for differentiability, giving 24 total measurements to calculate the limits for every colour in the RGB colour cube.

5.5 Implementation

Now that the necessary mathematical background has been presented, the implementation can be illustrated. This involves a description of the actual system, as well as design decisions that were made during its development.

5.5.1 System

The process of measuring the set of 24 RGB cube corner limit values (as described above, in Section 5.4), involves a pipe-and-filter approach. First, some basic tasks are completed to ensure the individual can easily discern between every pair of colours that define the edges of the RGB colour cube. This is done in order to identify and screen out individuals who are not able to differentiate between these corners. Differentiation between these corners is an integral component of the remainder of the system. If the individual successfully completes this initial task, the second stage is presented to them. In the second stage, the corner limit values are measured using a process described below. The results of the second stage need to be modified slightly through a linear extrapolation technique, which is the third stage. The extrapolated values from this stage are then fed to the predictor constructor, which uses these values to build the predictor, which is the final result of the system. This predictor is a piece of software that accepts any colour and calculates that colours differentiation limits. The predictor can also be given two colours, and it will determine whether the two colours are differentiable or not. This pipe-and-filter system is illustrated in Figure 5.2.

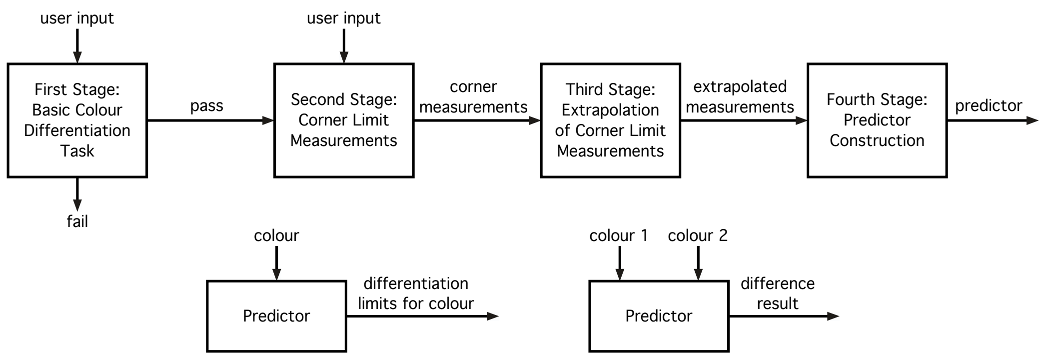

Figure 5.2: Overall pipe-and-filter design of the implementation for the model of colour differentiation.

Moving through this process is called *calibration* of the model. Through this process the general model of colour differentiation is calibrated to a specific environment and user (a context).

5.5.2 First Stage

In the first stage of the calibration process, a rough evaluation of each participant's colour differentiation abilities is performed. As the second stage (Section 5.5.3) measures differentiation at the endpoints of all external edges of the RGB cube, this initial stage was designed to ensure that the two endpoints along every external edge of the RGB cube are discernibly different. To see whether the endpoints of every external edge of the RGB cube are discernibly different, each possible pair of colours are shown to the participant and they respond whether the colours are different or not different using buttons at the top of the screen.

To accomplish this evaluation, the participant is presented with a rectangular field of circles, each approximately one `cm` in diameter, and set against a black background. The stage is divided into 12 individual trials. In each trial, about half of the circles are one colour, and the remainder are another colour. Using buttons at the top of the screen, the task is to identify whether the colours presented are different or not different (essentially whether they see two colours or one). When the participant selects one of the buttons at the top (either 'NOT different' or 'Different') the current trial finishes and the next trial is presented with new colours. The interface for this stage is shown in Figure 5.3.

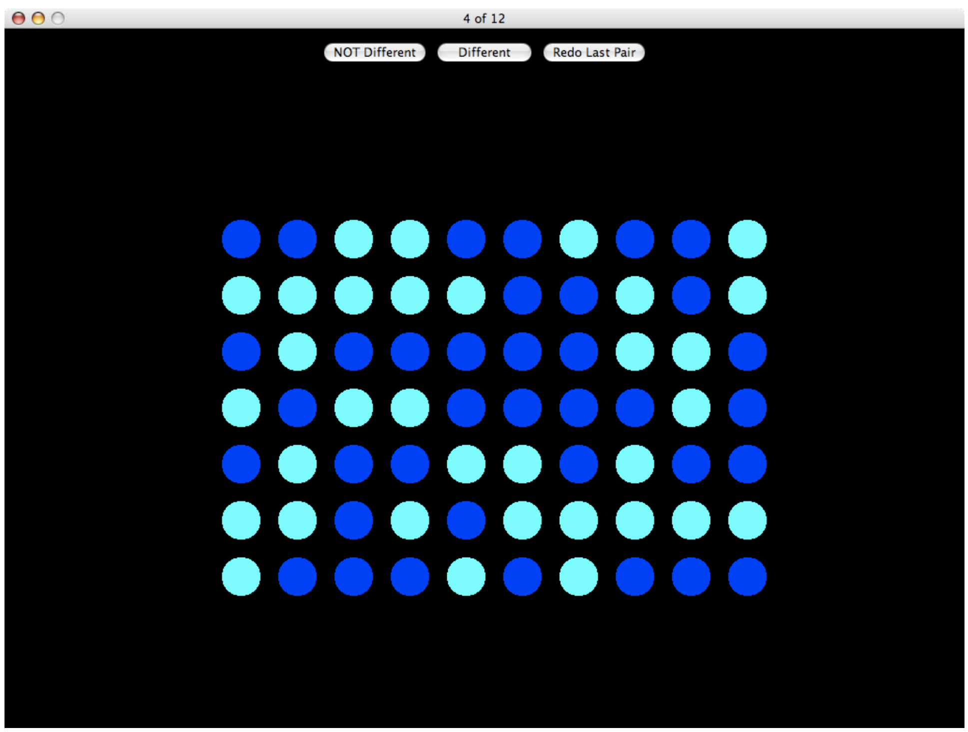

Figure 5.3: First stage of the calibration process.

The participant's response to each pair of colours is recorded. All of the pairs of colours presented to the participant are summarized in Table 5.1. If the participant responds that a particular pair of colours are not discernibly different, the calibration process is ended, as this response indicates that the participant will not be able to accurately complete the remainder of the calibration procedure.

First Colour	Second Colour	First Colour	Second Colour
Black	Blue	Green	Yellow
Black	Green	Red	Purple
Black	Red	Red	Yellow
Blue	Cyan	Cyan	White
Blue	Purple	Purple	White
Green	Cyan	Yellow	White

Table 5.1: Each pair of colours presented to the participant in the first stage of the calibration process.

5.5.3 Second Stage

In the second calibration stage, the participant is again presented with a rectangular field of circles (like the first stage), each approximately one `cm` in diameter, and set against a black background. The circles all begin as one colour, but approximately half of the circles can have their colour modified using a slider at the top of the screen. Using this slider, the participant's task is to identify the point at which the non-changeable colour and the slider-controlled colour become differentiable. When the participant identifies this point, they select a button labeled 'Proceed', the current slider value is recorded, and a new colour is presented to them. The participant is free to explore the entire range of possibilities in order to identify the point of interest. The interface for this stage is shown in Figure 5.4.

The slider controls the value of a specific channel for the modifiable colour. As an example, suppose the non-changeable colour is blue (0,0,255), and the channel of interest is the red channel. The non-changeable colour remains blue for the entire trial, but the slider is linked to the red channel value of the slider-controlled colour. The slider-controlled colour begins as blue (as the slider begins at '0'), but as the participant increases the value of the slider, the amount of red in the slider-controlled colour is increased too. As the slider moves, the colour displayed to the participant is updated in real-time to facilitate natural exploration of the range of the slider-controlled colour. In this case, the red channel can have any value between 0 and

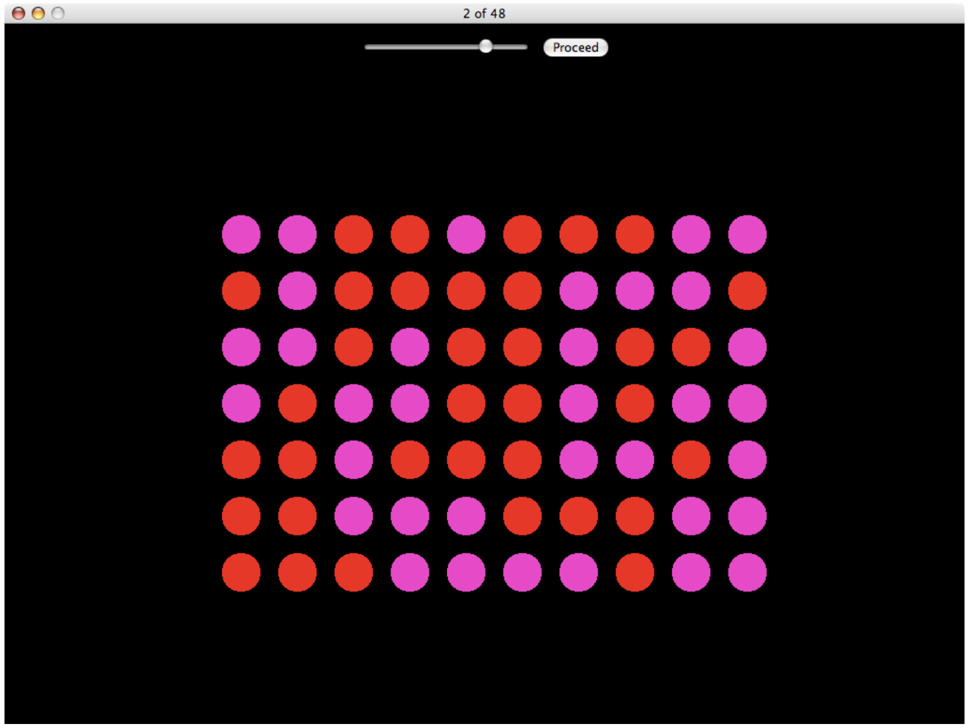

Figure 5.4: Second stage of the calibration process.

255 (inclusive), as it begins at 0 and can have a maximum value of 255. The remaining two channel values do not change. Suppose the participant selects the colour (13,0,255) as being just distinctly different from (0,0,255). When the participant presses the 'Proceed' button, the red channel value (13) is recorded as a red channel **upper** limit for the RGB colour cube corner (0,0,255).

As another example, suppose the non-changeable colour is white (255,255,255), and the channel of interest is the green channel. In this case both colours begin as white, and the slider is linked to the green channel value of the slider-controlled colour, but now the participant *decreases* the value of the green channel using the slider. Of course, the participant can still explore the entire range of possible slider-controlled colours (by moving the slider to different values), but as the green channel begins at 255, the participant will end up selecting a value less than 255 to define the point of differentiation. Suppose the participant chooses (255,214,255) as the point at which the two colours are just distinctly different. When the participant presses the 'Proceed' button, the green channel value (214) is recorded as the green channel

lower limit for the RGB colour cube corner (255,255,255).[3]

The colours presented to the participant are listed in Table 5.2. These are the colours represented by the corners of the RGB colour cube. For each colour, an upper or lower limit is measured for each channel, giving 24 trials[4]. Each trial is performed two times to give a total of 48 trials for the second stage of the calibration process.

Name	Red [0-255]	Green [0-255]	Blue [0-255]
Black	0	0	0
Blue	0	0	255
Green	0	255	0
Cyan	0	255	255
Red	255	0	0
Purple	255	0	255
Yellow	255	255	0
White	255	255	255

Table 5.2: Non-modifiable colours used in the second stage of the calibration process.

Exactly half of the data points collected will be upper differentiability limits (increasing channel values), and the other half will be lower differentiability limits (decreasing channel values). As the set of lower limits for all colours can be calculated from the set of upper limits for all colours (see Section 4.5), only the upper limits are desirable. Converting the half of the measurements that are lower limits to upper limits is the purpose of the third stage.

5.5.4 Third Stage

In the second stage of the calibration procedure (Section 5.5.3), two repetitions of 24 measurements are gathered from the participant. Within these 24 values are 12

[3] the lower limit is actually $255 - 214 = 41$.

[4] eight corner colours $\times$ three channels per colour

pairs of values, one pair for each of the outside edges of the RGB colour cube. As described in Section 5.3, the endpoints for an edge are those limits for the channel that changes as one moves along the edge. One of the limits will be an upper limit (because the channel of interest is equal to zero at one end of the edge), and the other limit will be a lower limit (because the channel of interest is equal to 255 at the other end of the edge). Each of these edges, and their corresponding upper limit and lower limit corners are given in Table 5.3. Using these two values and the values of the channel of interest at each of these points, the upper limit can be extrapolated to the '255' end of the edge.

Testing Channel	Colour Edge	Upper Limit Corner	Lower Limit Corner
Red	Black → Red	Black	Red
Red	Green → Yellow	Green	Yellow
Red	Blue → Purple	Blue	Purple
Red	Cyan → White	Cyan	White
Green	Black → Green	Black	Green
Green	Red → Yellow	Red	Yellow
Green	Blue → Cyan	Blue	Cyan
Green	Purple → White	Purple	White
Blue	Black → Blue	Black	Blue
Blue	Red → Purple	Red	Purple
Blue	Green → Cyan	Green	Cyan
Blue	Yellow → White	Yellow	White

Table 5.3: Testing channel, endpoint colours, lower limit corners, and upper limit corners for RGB cube outside edges.

As an example, suppose the edge that goes from green to cyan is under consideration. As one moves along the edge from green to cyan, the blue channel value steadily increases. Because of this, the blue channel is the channel of interest (for which the upper limit function is desired). At the green corner (0,255,0), blue is 0, and therefore the value measured in the second stage for the blue channel is an upper limit. Suppose this value is 34 (i.e., the participant stated (0,255,0) and (0,255,34) are distinctly different colours). At the cyan corner, the blue channel is 255, and therefore the measurement made here is a lower limit. Suppose this value is 76 (i.e., the participant stated that (0,255,255) and (0,255,(255-76 = 179)) are distinctly different colours).

To convert this lower limit to an upper limit for the (0,255,255) corner, a linear extrapolation needs to be performed. The key to performing this extrapolation lies in the fact that because 76 is a lower limit for the blue channel at (0,255,255), 76 can also serve at the upper blue limit for (0,255,179) (i.e., the participant has essentially stated that at (0,255,179), the upper blue channel limit is 76). The linearity assumption can then be used to extrapolate the linear function defined by these two points to determine a upper blue channel limit for (0,255,255). This is illustrated in Figure 5.5, where the desired upper limit is marked with $L_?$, and is equal to:

$$L_? = (76 - 34)/179 * 255 + 34 = 93.8$$

So the desired upper blue channel limit for (0,255,255) is 93.8.

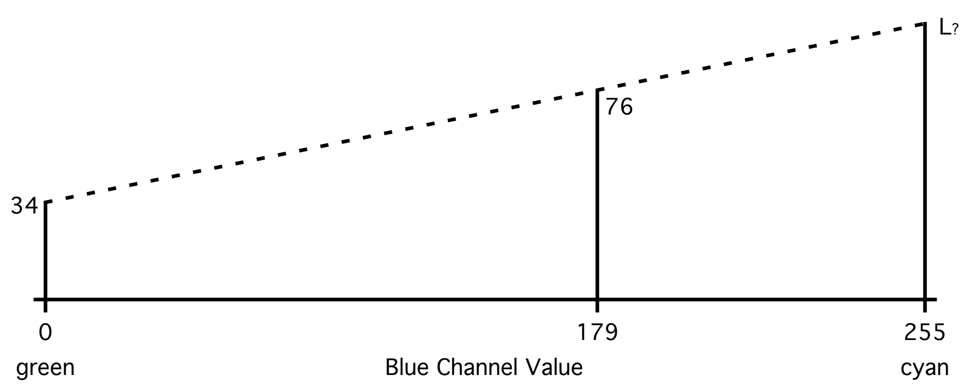

Figure 5.5: Extrapolation process to convert lower limits to upper limits.

These extrapolations are necessary to perform, as they greatly reduce the complexity of calculating the slope and intercept for the linear functions of the model. Because this extrapolation process sets the two known limits for every function to precise locations (namely the RGB colour cube corners), the arithmetic of these functions is simplified.

This process of extrapolation is performed outside of the software using a spreadsheet application. This was to allow rapid refinement of the process, as well as clear verification of the results. Of course, this extrapolation process could be integrated into the software system with minimal difficulty.

5.5.5 Fourth Stage

The sole purpose of the fourth stage of the calibration process is to produce the final colour differentiation predictor - the piece of software that can be used to make estimations of the per-channel differentiation limits for a given colour. This is simply called the predictor.

To construct the predictor, each of the 24 limit values (the 12 original upper limit values and the 12 extrapolated upper limit values from Stage 3), are used to generate the slope and intercept for each of the linear functions that are defined along the edges of the RGB colour cube. As these limits are for pre-defined locations of the RGB colour cube (the corners), none of the locations need to be supplied to the predictor constructor. The fact that the limits provided are at the corners allows simple calculation of the intercepts and slopes. For each function, the intercept is simply the upper limit value for the corner at which the testing channel value is zero (e.g., when testing on the red channel for the edge (green $\rightarrow$ yellow), the upper red channel limit for the green corner). To find the slope, some calculation is required, but is simply the extrapolated upper limit value minus the original upper limit value divided by 255. This is because the 'rise' (ΔY) is just the difference between the two limits, and the 'run' (ΔX) is the total change of the channel of interest over the entire edge (which starts at 0 and goes to 255 courtesy the extrapolation process of Stage 3). Slope is $\Delta Y / \Delta X$.

Once the predictor is constructed, it provides one main service. This is to accept any colour (specified as a triple of values in the range [0-255]), and to generate the upper and lower limits for each channel for that colour. These limits can then be used to make predictions about the perceptual difference between this colour and any other colour.

To generate these limits, the process described in Section 5.1 for linearly interpolating into a 3D volume is followed. First, four functions are used to calculate four points which define a plane in the volume. This plane is defined by two new functions, and these are used to calculate a single function. Finally this function is used to calculate a single point. This point represents the upper differential limit for a given channel for the supplied colour. This process is repeated (with different parameters) for the remaining channels.

The reason this piece of software is called a 'predictor' is twofold. First, because of the drastically reduced number of samples that are used to generate this model of colour differentiation, it can best be described as making a prediction, based on the truthfulness of the linearity assumption. If there are some cases in which the linearity assumption is incorrect, then the values supplied by this software may be inaccurate. The second reason this software is called a 'predictor' is that it uses the functionality described above to supply an additional service. This service accepts two colours, and using the definition of colour differentiability (given in Section 4.1.2), and the generation of limits for one of the colours, makes a binary *prediction* as to whether the other colour is perceptibly different from the first.

Chapter 6

Evaluating the Assumption of Linearity

A baseline study was carried out in order to answer basic questions about the feasibility of the model of colour differentiation presented in Chapters 4 and 5. There were a set of unanswered questions about how well the assumption of linearity would work as a basis for the implementation of the general model. The assumption of linearity allowed a substantial reduction in the number of limit measurements required to construct a context-specific model of colour differentiation, hence it was very important to assess the validity of this assumption.

6.1 Goals

In the model presented in Chapter 4, the discernibility of two colours is based on individual channel differences between the two colours. The first question that arises from this definition is how the starting value of a channel affects differentiability. This definition also needs to be considered in terms of variation in the remaining channels. This results in the following questions:

1. With remaining channels held constant, how does the starting value of a channel influence differentiability along that channel?

2. How does the starting value of a channel influence differentiability along that channel when the remaining channel values are variable?

This study was aimed at examining the relationship between differentiability and channel values for a colour before the entire modeling system was implemented and studied further. The intention was to use the results of this study to confirm the assumption of linearity upon which the general model of colour differentiation is implemented.

6.2 Study Methods

6.2.1 Participants

Six participants, all male, were recruited from the University of Saskatchewan. All were students or graduates from technical degree programs, and ranged in age from 24 - 38 years (mean age: 32.2 years). All had normal (or corrected-to-normal) visual acuity, and had no previous diagnosis or history of colour blindness. All participants were regular computer users (more than ten hours per week).

6.2.2 Apparatus

The study used a custom-built Java application. All participants performed the study using a laptop paired with an external CRT monitor running at a resolution of 1024 x 768. A CRT monitor was used because CRTs are less prone to angular colour variation which is a frequent issue with LCD monitors (and a potential source of error in this study). The study was performed in full-screen mode on the external monitor. Participants interacted with the program using an external mouse. Keyboard input was not necessary. See Figure 6.1 for a screenshot of the system interface.

6.2.3 Tasks

The task was to identify the colour differentiability limits for colours at several points in the RGB colour cube.

As can be seen in Figure 6.1, the user was presented with a rectangular field of circles, each approximately one `cm` in diameter. About half of the circles were one colour, and the remainder were another colour. One of these halves remained the same colour for an entire trial, but the other half could be manipulated by the participant. The colour that remained consistent will be called the non-changing

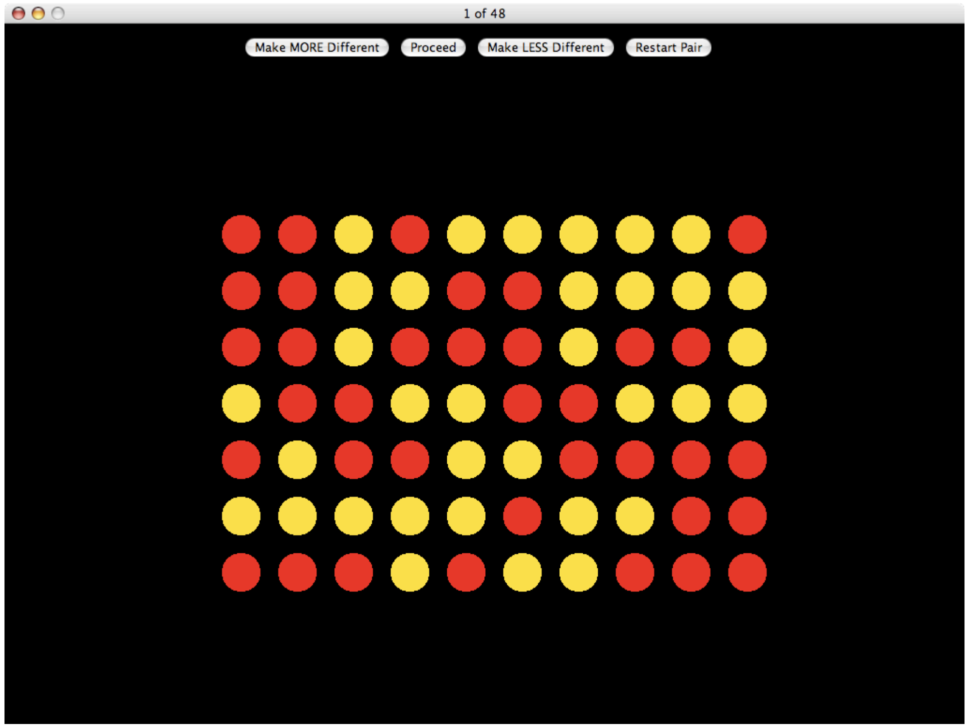

Figure 6.1: System used for colour discrimination tasks.

colour. The colour that was participant manipulated will be called the manipulated colour.

The participant's task was to use buttons at the top of the window to adjust the manipulated colour until it was just differentiable from the non-changing colour. The task was divided into a number of individual trials. In each trial, one button (labeled 'Make MORE Different') would be used to make the colours displayed more different. Another button (labeled 'Make LESS Different') would be used to make the displayed colours less different. These buttons controlled the amount of single channel difference between the non-changing colour and the manipulated colour. The participant used these buttons to adjust the difference between the two colours, and when the point at which they considered the colours to be just differentiable was reached, the participant clicked on a button labeled 'Proceed'. If the differentiability point could not be found (for reasons outlined in Section 6.2.5), the participant could use the 'Restart Pair' button to begin the current colour pair again. There were 48 unique pairs of colours (each comprising a single trial) that the participant worked through, ordered randomly.

This task was not a performance task, but a judgement task. The participant was asked to determine the point at which the two colours were just differentiable. This is an inherently judgmental task, as this point does not represent an absolute transition from not differentiable to differentiable. This is because this point resides in a region of uncertainty. At the lower end of this region, the colours will not be differentiable at all. At the higher end of this region, the colours are unquestionably differentiable. Between these two extremes, the differentiability of the two colours is not precise. Because of this lack of precision, it was decided that the task of identifying the differentiable point could be offloaded onto the participant by having them perform a judgement task.

6.2.4 Procedure

Each study session involved a single participant at a time. The session procedure had five steps:

1. **Pre-Study Questionnaire**: This questionnaire was used to gather personal information about the participant. Personal information included age, gender, occupation, education level, computer experience (hours per week), electronic gaming experience (hours per week), information visualization experience (visualizations and tools used), corrected visual acuity (excellent, good, fair, poor), previous diagnosis of colour blindness, and family history of colour blindness. A sample pre-study questionnaire is given in Appendix A.

2. **Delivery of Verbal Instructions**: The participant was informed that the tasks they were about to perform were not an evaluation of their abilities, in that they could not fail at the task. The participant was informed that time was not being recorded, and the sole purpose of the study was to measure their response to a series of tasks, and that they could quit the study at any time. A verbal description of the task was then given, including a specific description of what 'just differentiable' meant (the minimum point at which

the colours were differentiable at a glance). The participant was asked to attempt to be consistent in their judgement of the 'just differentiable' point (see Section 6.2.5), as the task was a judgement task, not a performance task. Last, the user was informed of the approximate time to complete the task (about 30 minutes).

3. **Program Instructions**: At this point the program was started, and the user was asked to read the written instructions for the study. These instructions are shown in Figure 6.2.

4. **Training**: The participant then performed three trials of the study to ensure they were confident in their understanding of the expectations from the task. The program was then stopped and restarted. No results were recorded from this stage.

5. **Experimental Tasks**: The participant then performed 48 trials, taking intermittent breaks to rest their eyes (every 12 trials).

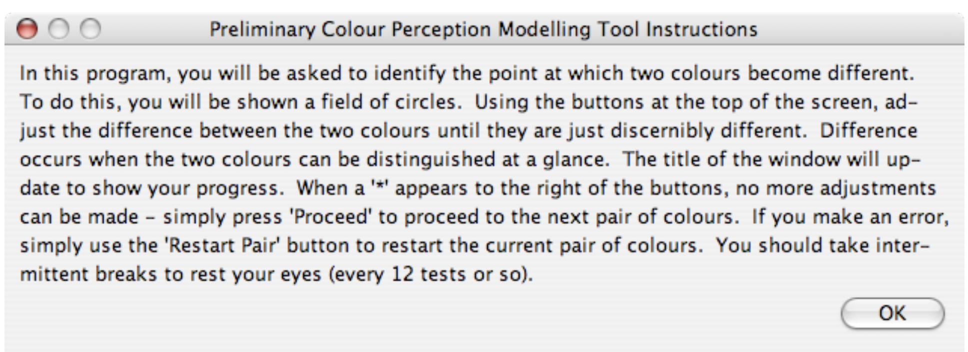

Figure 6.2: Instructions given at beginning of study.

6.2.5 Study Design

As this was a baseline study, data was collected to determine the relationships that exist between colour differentiation values in a digital colour environment. This data was analyzed to see whether the linear assumption was a reasonable hypothesis.

In this study, the point at which colours become differentiable was examined. The task involved the participants finding the point at which two colours became just differentiable. To determine differentiability, the definition given in Section 4.1 was used. This definition states that two colours can be differentiated when their respective channel values differ sufficiently. This study aimed to give a quantitative value to the term 'sufficiently'. The units of this value were RGB channel intensity values. This value is defined as a *limit* in Chapter 4. To measure the point at which two colours become differentiable, participants were asked to vary the single channel difference between two colours until they were just differentiable. This value will be referred to as the 'differentiability value'.

The independent variable in the study was the non-changing colour (specifically the RGB channel values for the non-changing colour). The dependent variable was the just differentiable point indicated by the participant. The possible values for each channel of the non-changing colour were 0, 85, 170, or 255. As this study tested for differentiability via a single channel increase (upper limit), the channel being measured (called the *testing channel*) was never given the value 255, as no increase is possible from this value. This resulted in the following number of conditions:

$$\begin{array}{rl} & \text{3 channels} \\ \times & \text{3 values on testing channel} \\ \times & \text{4 values on first remaining channel} \\ \times & \text{4 values on second remaining channel} \\ \hline & \text{144 total conditions} \end{array}$$

To measure the differentiability value, a binary search approach was used. When a pair of colours was first presented, the differentiability value was set to $(255 - T_{NCC})/2$, where T_{NCC} is the value of the testing channel ($T \in \{R, G, B\}$) for the non-changing colour (NCC). If the participant selected 'Make MORE Similar', the differentiability value was decreased to $(255 - T_{NCC})/4$, and all possible values of the differentiability value greater than $(255 - T_{NCC})/2$ were discarded. If the participant selected the 'Make LESS Similar' button, the differentiability value was increased to

$3 \times (255 - T_{NCC})/4$, and all possible values of the differentiability value less than $(255 - T_{NCC})/2$ were discarded. The differentiability value was increased or decreased in a binary fashion until it could only be one value (as binary search repeatedly eliminates the upper half or lower half of the possible values to search, it eventually reaches only one possible value). This is the differentiability value that was recorded. Using this binary approach allowed the participant to quickly proceed to a value (in less than eight steps), and the speed of the process was why it was selected.

Another benefit of the binary search approach was that it provided a single differentiability value to be recorded. A difficulty that was encountered while designing this experiment was that there was often an area of uncertainty, in which the available differentiability values allowed some discernment to occur between the colours, but not enough for the participant to declare the colours just differentiable. Instead of developing some approach to deal with this lack of precision, the binary search approach allowed the participant to select the point at which they deemed the colours distinct. This is referred to as a *judgement task* for the user, and worked well to handle the situation.

Although the binary search approach provided a fast and precise way to find differentiability values, it did seem to impose a load on the participant. Some participants found the task quite difficult. To help alleviate this load, it was decided to limit the number of trials per user to 48. This was accomplished by having only one channel of testing per participant. Participants one and four performed the tasks on the red channel exclusively, participants two and five performed the tasks on the green channel exclusively, and participants three and six performed the tasks on the blue channel exclusively. This seemed to be a good number of trials as the participants could generally be finished in about half an hour.

Only six participants performed this study. It is not believed that this is a representative sample of the human race, but as this was a feasibility analysis, this number of participants was deemed sufficient.

6.3 Results

Overall results of this study are presented first visually. After this, both questions asked in Section 6.1 are examined in light of the results.

6.3.1 General Results

As each participant performed the study on exclusively one channel, three images are presented, one for each channel. Each image is a *small multiples* visualization [62], in which the scale is reduced to get an overall visual impression of the data. Figure 6.3 presents one participant's results for red channel testing, Figure 6.4 presents one participant's results for green channel testing, and Figure 6.5 presents one participant's results for blue channel testing. Each figure is a composition of 16 individual charts. Each individual chart plots the same data - the starting value for the channel of interest (e.g., Red in Figure 6.3) versus the differentiation value, but the value for the remaining channels varies (e.g., Green and Blue in Figure 6.3) between charts. The values for the remaining channels can be found in the top horizontal and left vertical regions. For each individual chart, the independent variable (x-axis) is the starting value for the testing channel (one of 0, 75, or 180), and the dependent variable (y-axis) is the value that needed to be added in order for the participant to distinguish along the testing channel (the differentiation value, presented above in Section 6.2.5).

6.3.2 Specific Results

In Section 6.1, questions were posed that this study was designed to answer. Here each question is examined individually and results are presented to answer each question.

Question #1

The first question posed was:

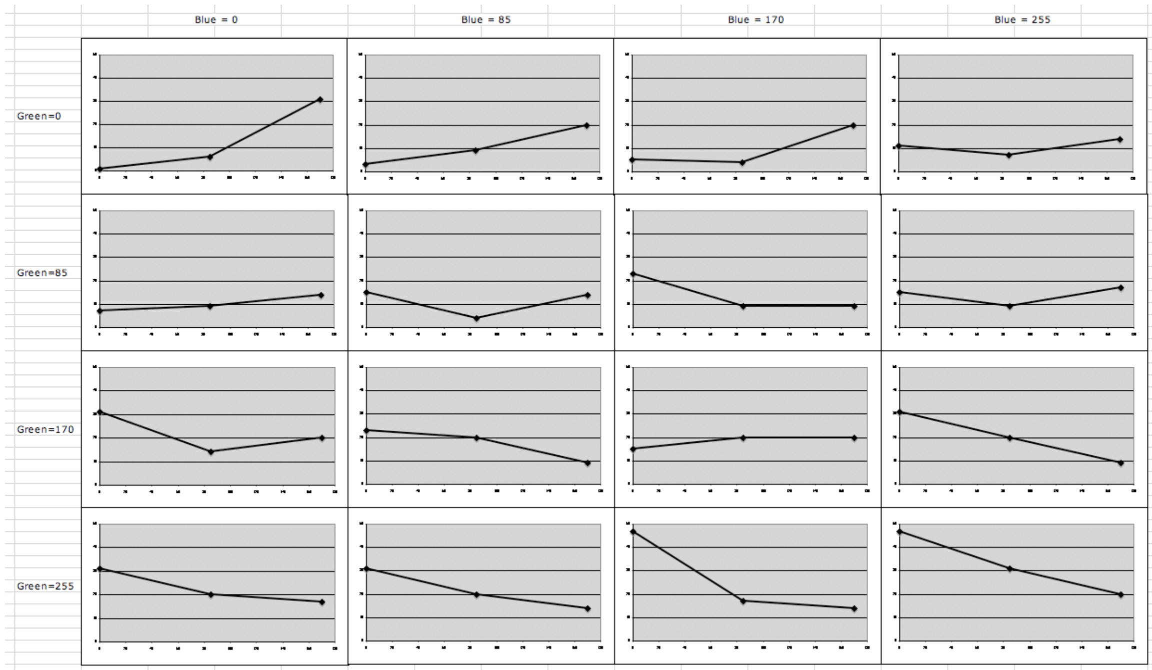

Figure 6.3: Red channel study results for a single participant presented using small multiples. It can be seen that red channel differentiability increases with channel start value for low green channel values, but decreases with the channel start value for high green channel values. Blue channel values do not appear to affect red channel differentiability, as the individual charts are fairly consistent across columns.

> With remaining channels held constant, how does the starting value of a channel influence discernibility along that channel?

Results presented in Figures 6.3-6.5 suggest that some relationship exists between starting channel value and differentiation value. It is hypothesized that this relationship is linear, and here present data in terms of this hypothesis.

A *linear relationship* can be used to describe a relationship between two variables x and y such that $y = mx+b$. As a plotted function of x versus y, a linear relationship produces a straight line plot. For this line, m is the slope and b is the y-intercept (the value of y when $x = 0$).

To measure how well the data from this study supports the linear relationship hypothesis, the coefficient of determination (Pearson's correlation coefficient squared or R^2), was calculated. The results of this calculation are given in Table 6.1.

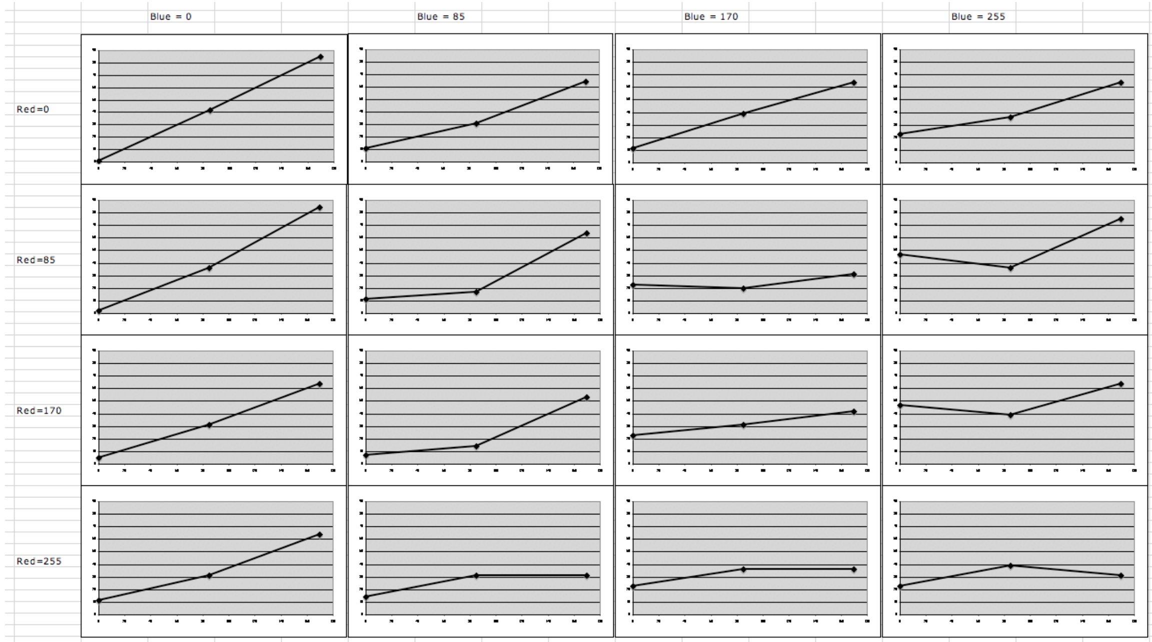

Figure 6.4: Green channel study results for a single participant presented using small multiples. It can be seen that green channel differentiability increases with channel start value for most cases, except when red channel is high and blue channel is non-zero.

As can be seen from the R^2 values presented in Table 6.1, there was significant variation within subjects (minimum R^2 value was near zero for each participant, and maximum R^2 value was near one for each participant), but less variation between subjects (mean [0.7-0.9] and median [0.8-1.0] R^2 values are fairly consistent between subjects). Each participant scored at least one low and at least one high R^2 value. The mean R^2 seems to indicate that there are more strong correlations than weak correlations (as the mean is considerably above 0.5). The median R^2 values support this idea as well.

What results in some of the relationships having high R^2, and some having low R^2? One explanation is that Pearson's correlation coefficient does not test how well the data fits a linear function, but how well the independent and dependent variables are linearly coordinated. As such, when the dependent variable does not depend on the independent (and remains constant), a flat line plot results, but the R^2 value is zero or undefined. This situation is simply a special case of the linear function in which the slope, m, is zero. These zero correlations do not matter in the larger

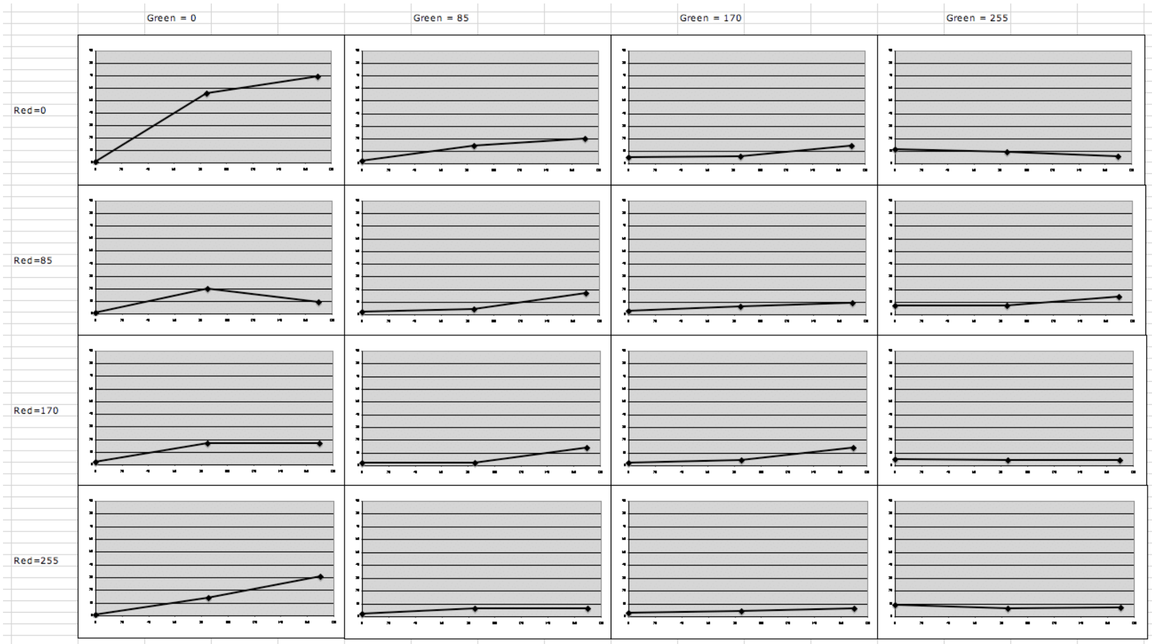

Figure 6.5: Blue channel study results for a single participant presented using small multiples. It can be seen that blue channel differentiability does not vary much at all, except when the red and green channel values are zero.

sense, as a linear function with slope of zero still describes the relationship between independent and dependent variables well.

Instead of using Pearson's correlation coefficient, it was determined that another evaluation of the results was needed. To accomplish this, the *least squares* method was used to generate values m and b for a linear function that most closely approximates the relationship between independent and dependent variables. The amount of error between this linear function and the actual data results was then measured. This was done by calculating the y-axis difference between the function and the actual data for each of the independent variable values (0, 85, 170) to give three values. This is illustrated in Figure 6.6. The actual data points are labeled P_1, P_2, and P_3, and the best-fit linear line (linear regression line) is labeled R. The lengths of each arrow A, B, and C, represent the error score described here. As the best-fit line lies between the actual data points, some of the arrow lengths are negative, and some are positive. These error values were used as a raw evaluation score, and have implications for the remainder of this research (as described in Section 6.4). This

Participant	Channel	Min R^2	Max R^2	Mean R^2	Median R^2
1	Red	0.0	1.0	0.7	0.8
2	Green	0.3	1.0	0.8	0.9
3	Blue	0.2	1.0	0.8	0.8
4	Red	0.0	1.0	0.7	0.8
5	Green	0.1	1.0	0.7	0.9
6	Blue	0.2	1.0	0.9	1.0

Table 6.1: Correlation coefficient values for this study. Non-testing channels held constant.

error score will be referred to simply as the *difference*. The units for difference are in channel intensities, which range in value from 0 - 255. A summary of the results from the difference data analysis are presented in Table 6.2.

Participant	Channel	Min	Max	Median	Mean	Std Dev
1	Red	-9	4.5	0.8	< 0.0001	3.3
2	Green	-16.7	8.3	0.8	< 0.0001	5.4
3	Blue	-5	10	0.2	< 0.0001	2.5
4	Red	-33.7	23.3	1.5	< 0.0001	11.6
5	Green	-5.7	5	0.7	< 0.0001	2.4
6	Blue	-7	14	-0.3	< 0.0001	4.7
Overall:		-33.7	23.3	0.5	< 0.0001	5.8

Table 6.2: Characteristics of the difference values between best-fit linear function and participant data. 48 data points for each participant gives 288 total results overall. Non-testing channels held constant.

As the mean and median values are quite similar, and the minimum and maximum values are somewhat balanced, it appears that this data may be normally distributed around the mean (~ 0). A histogram of the overall results is shown in Figure 6.7.

By aggregating the difference values for all participants, an overall mean of ~ 0 and standard deviation of 5.8 are found. A mean of ~ 0 is not surprising, given the method of *least squares*. This method looks to minimize the sum of the squared

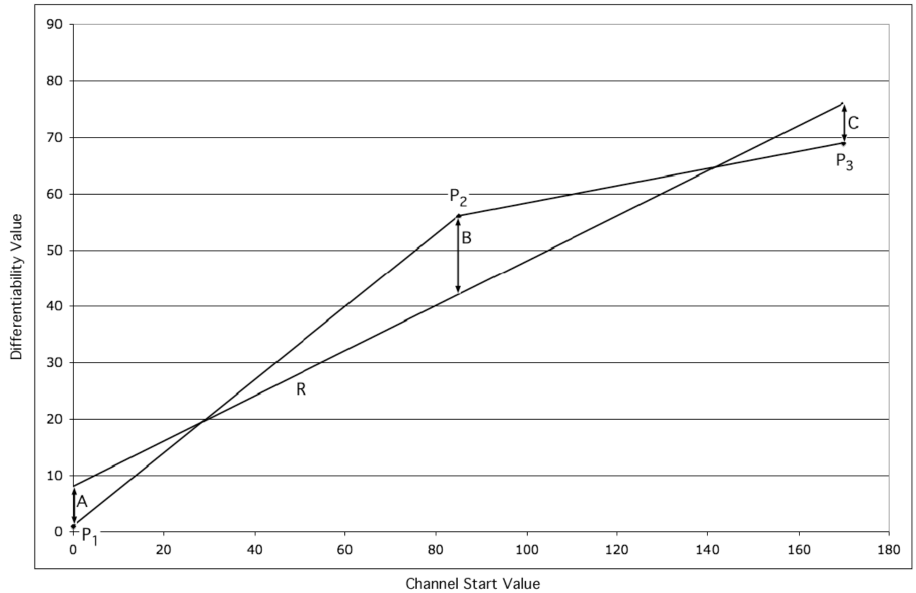

Figure 6.6: Visual explanation for Difference. The length of arrows A, B, and C are the difference (+/-) values for this chart.

difference values the have been calculated here. The minimization of this sum will result in the values being generally balanced on both sides of the linear function. This will result in a mean of nearly 0, because the error across all datapoints will be balanced over the linear function. The standard deviation is a very useful value, as it gives the amount of error that can be expected by using the best-fit linear function to make predictions about colour differentiability. This will be discussed further in Chapter 8.

Question #2

The second question posed was:

> How does the starting value of a channel influence discernibility along that channel when the remaining channel values are variable?

To answer this question, Figures 6.3-6.5 can be examined in a slightly different way than how they were examined to answer the first question. In this case, the same independent variable value will be considered, but across columns and rows of the small multiples visualization. Although this could be visualized directly by drawing lines in these figures joining values between individual charts, it can be

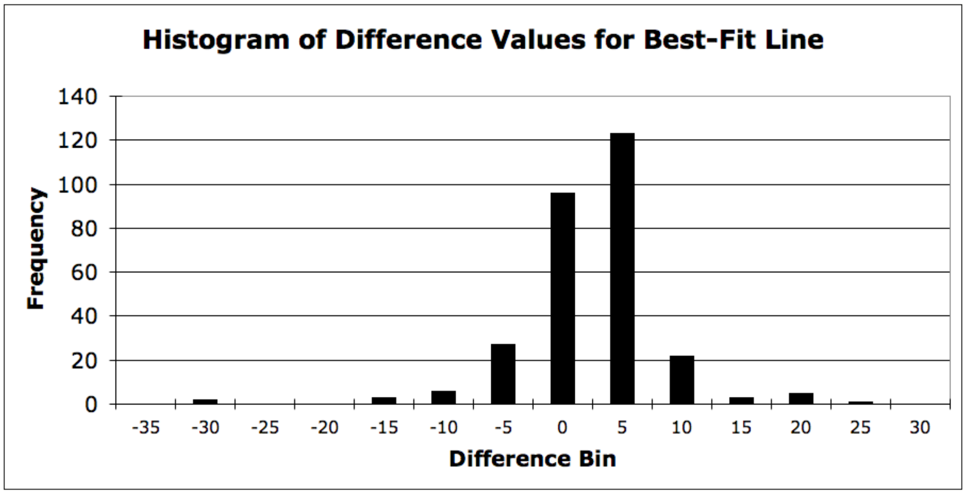

Figure 6.7: Histogram distribution for the difference values between best-fit linear function and participant data. Non-testing channels held constant.

presented in a clearer fashion. By modifying the major axes of the small multiples visualization, the same information can be visualized while maintaining the atomic nature of each chart. To do this, one of the axes (in this case the vertical axis on the left side of the visualization) is set to the possible values of the testing channel. The horizontal (top) axis is arbitrarily chosen to be one of the remaining channels, and the other remaining channel is set as the independent variable in each chart. This has the affect of visualizing how the differentiability measure for the testing channel varies according to the variance of one of the non-testing channels. As there are two non-testing channels, this results in two possible visualizations for each participant.

As an example, suppose that a participant determined differentiability values for the blue channel. The relationship between these differentiability values and variance along the non-testing channels (red and green in this instance) is being explored. This is performed by plotting the variation in differentiability values against the value of the non-testing channel, once for each of the two non-testing channels. This results in a small multiples visualization in which the testing channel start value now moves outside of the individual chart (to the vertical axis in this case). The non-testing channel of interest moves to be the independent variable for the individual charts, and the remaining non-testing channel is placed along the horizontal axis of the visualization. This results in two visualizations per participant, as there are two non-testing channels. Figure 6.8 contains an example of this visualization, in which blue is the testing channel and green is the non-testing channel of interest. This visualization shows how blue channel differentiability varies in relation to green channel values. Figure 6.9 contains the counterpart to Figure 6.8. In this visualization, blue is still the testing channel, but now red is the non-testing channel of interest. This visualization shows how blue channel differentiability varies in relation to red channel values.

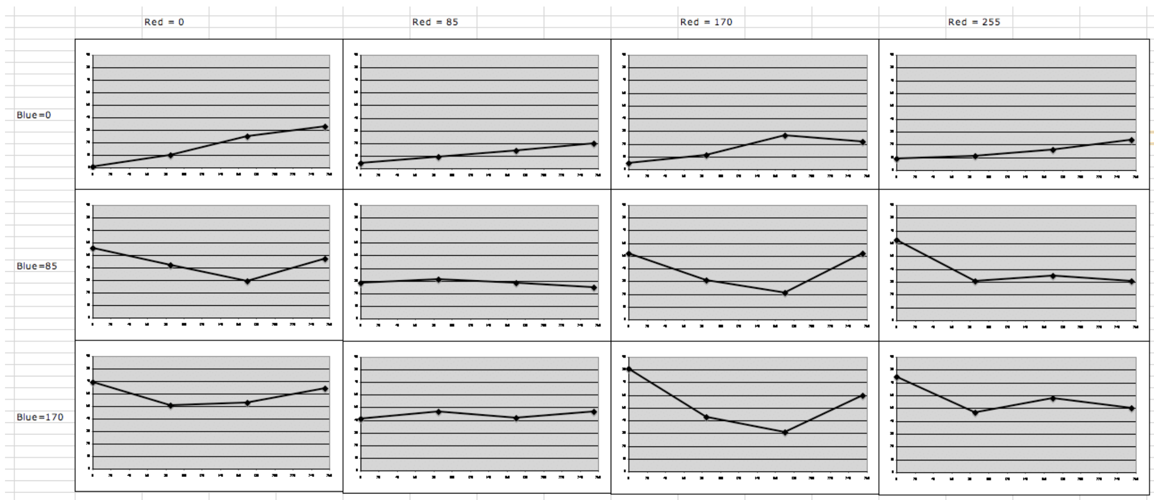

Figure 6.8: Blue channel differentiability related to green channel variation. It can be seen that when blue = 0, differentiability along the blue channel increases in a consistent manner with the green channel value, independent of red channel value. When the blue channel is non-zero, however, this relationship breaks down, particularly for the cases where red = 170.

As Pearson's correlation coefficient squared (R^2) produced similar variations as those found when answering the first question above, it was decided that only overall summaries of R^2 values would be presented (Table 6.3). These summary values were aggregated over all participants. As can be seen, the average and median R^2 values for the aggregated data set are not very different, and the overall range spans the total possible range for R^2 values [0-1].

The results of difference analysis are given in Table 6.4. As there are two non-testing (independent) channels per testing channel per participant, the data gathered from each participant is presented, one for each of the non-testing channels. Aggre-

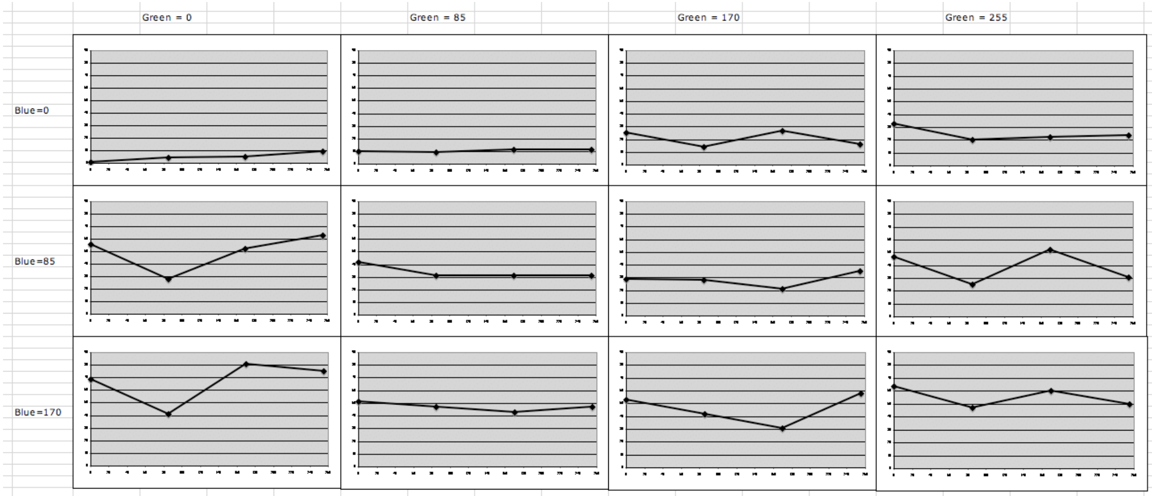

Figure 6.9: Blue channel differentiability related to red channel variation. It can be seen that blue channel differentiability does not depend substantially on red channel value for most cases. Exceptions to this occur when blue is non-zero, and green = 0 and green = 255, where the relationship is chaotic.

Min R^2	Max R^2	Mean R^2	Med R^2
0	1.0	0.5	0.4

Table 6.3: Aggregated results showing Pearson's correlation coefficient squared (R^2) information for all participants. This shows that there was considerable variation in R^2 values over the entire data set.

gate information is also presented to provide an overall feel for the data.

6.4 Analysis and Implications

The results presented above are now analyzed to evaluate the feasibility of the assumption of linearity. Some implications of this study will also be presented.

6.4.1 Summary of Results

In Section 6.1, two questions were put forward. The purpose of this study was to answer these two questions. These questions were:

1. With remaining channels held constant, how does the starting value of a chan-

Part	Independent	Channel	Min	Max	Median	Mean	St Dev
1	Blue	Red	-10	7.2	0.7	< 0.0001	3.5
1	Green	Red	-13	7.8	0.5	< 0.0001	4.2
2	Blue	Green	-30	20.5	0.4	< 0.0001	8.1
2	Red	Green	-16	12	0.7	< 0.0001	6.8
3	Green	Blue	-10	7.7	-0.1	< 0.0001	3.6
3	Red	Blue	-19	15	0	< 0.0001	4.4
4	Blue	Red	-46	30.6	0.4	< 0.0001	15.1
4	Green	Red	-34	44.2	1.5	< 0.0001	14.9
5	Blue	Green	-11	10.8	0.3	< 0.0001	5.6
5	Red	Green	-9	7.4	0	< 0.0001	3.3
6	Green	Blue	-19	17.5	0.1	< 0.0001	8.1
6	Red	Blue	-23	14.3	0.4	< 0.0001	7.7
Overall:			-46.2	44.2	0.3	< 0.0001	8.0

Table 6.4: Difference results showing how variance in the independent (non-testing) channel effects differentiability in the testing channel (named simply 'channel' in this table). 48 data points for each participant gives 576 total results overall.

nel influence discernibility along that channel?

2. How does the starting value of a channel influence discernibility along that channel when the remaining channel values are variable?

The results presented in Section 6.3 present data that can be used to answer these questions. In particular, the concept of *difference* was introduced as a metric to evaluate the hypothesis of a linear function describing the relationship between channel values and the differentiability of a particular channel. The aggregate results are provided again here for ease of reference in Table 6.5.

Analysis #	Min	Max	Median	Mean	Standard Deviation
1	-33.7	23.3	0.5	< 0.0001	5.8
2	-46.2	44.2	0.3	< 0.0001	8.0

Table 6.5: Mean and standard deviation information for aggregate data. Analysis #1 data can be used to answer Question #1, and analysis #2 data can be used to answer Question #2.

6.4.2 Implications of Results

The values presented in Table 6.5 indicate that a linear function *can* be used to describe the differentiability of a channel when the remaining channels are held constant. Although a best-fit linear function fits the data well, it does not model the data perfectly. As a result, some errors are expected, and these can be calculated using the standard deviations for the aggregate results.

The mean is ~ 0 in all cases analyzed. This is due to the nature of the *least squares* method used to determine the best-fit line (as outlined above in Section 6.3.2). This mean can be used with the standard deviation (SD) values to determine the range of realistic results using the best-fit linear function for prediction of colour differentiation. The range from (-2 $\times$ SD) to (2 $\times$ SD) will give a 95% coverage of the distribution. For the first analysis, the SD is equal to 5.8, giving a range of [-11.7, 11.7]. The second analysis has an SD of 8.0, giving a range of [-16.1, 16.1].

In terms of predicting colour differentiation, these ranges allow some confidence in the predictor described in Section 5.5. to be developed. If the predictor uses a linear function that is determined using the *least squares* method, then it can be expected that 95% of the predictions made using this predictor will be within these ranges.

There is some difficulty because two ranges are available, when a single range would be much more useful for generating a confidence for a predictor. To generate a single range, an aggregate analysis over *every* distance metric measured above can be performed. The results of this analysis are presented in Table 6.6. It can be seen that the mean is still ~ 0, but now the standard deviations have merged to a final value of 7.4.

Min	Max	Median	Mean	Standard Deviation
-46.2	44.2	0.3	< 0.0001	7.4

Table 6.6: Mean and standard deviation information for entire data set. A total of 864 data points are considered here, comprising all the values used in the first and second analyses.

Using these combined results, a final 95% percent coverage range can be calculated. With a mean of 0, and SD of 7.4, the range [-14.7, 14.7] is found. This range essentially combines the results from both of the analyses performed above to get a total confidence range for a predictor that uses a best-fit linear function. What this means is that when a linear function (determined using the *least-squares* method) is used to predict upper and lower differential limits for a given colour, the values given will be within $\pm$ 14.7 of the true value. Of course, if another approximation to the linear function is used instead of the *least-squares* method, additional uncertainty may be introduced. As this is the case with the implementation of this predictor described in Section 5.5.1, this confidence range serves as a rough approximation and nothing more.

6.4.3 Critical Reflections

In this section, some factors that potentially influenced the outcome of this study are presented and discussed. Two main factors have been identified, experimental design and sample size.

Experimental Design

Four of the six participants complained about the binary search approach used in this experiment (described in Section 6.2.5). With each press of the 'Make LESS Different' and 'Make MORE Different' buttons, the magnitude of the change would decrease. This led participants to feel as though they were being 'punished', in that they could not explore the possible answers with sufficient freedom, and were too restricted while performing the exercise. Ultimately, these participants explored the possible answers by exploring a path, and then restarting the trial. This was repeated until the participant found the point at which the colours became 'just differentiable'.

Although no solution was implemented for this problem during the study, participants generally agreed that the use of a slider instead of the binary search approach would still be fast, and allow the participant to thoroughly explore the possible solutions before making their selection. The use of a slider was implemented for the following study, presented in Chapter 7.

Sample Size

When answering Question #1, three data measurements were used to calculate the best-fit linear function, as well as to establish difference points against this function (each chart in Figures 6.3-6.5). It can be argued that fitting a linear function to three points is dangerous, considering that a linear function can always be perfectly fit to two distinct points. Although this weakness is recognized, this study was performed to establish whether a linear function would work well for the purposes of predicting colour differentiability, and this linear function will need to be constructed from very few points (two in the present implementation) to help prevent a time-consuming calibration procedure. Therefore, considering against more points, although desirable, will not be pursued further.

The above argument can also be made regarding the four data measurements used to calculate best-fit linear functions in order to answer Question #2. Again, more points would allow greater refinement of the best-fit linear function, and produce more samples of the difference metric, but are again not practical from the predictor calibration standpoint presented above.

An additional consideration is the experimental design of this study. As each participant's data was uniquely dependent upon several uncontrolled factors (such as colour perception abilities, lighting conditions, etc.), there is no way to gather more than three data points for Question #1 or four for Question #2 than to increase the number of trials per participant. As a result of the difficulties imposed by the binary search approach used in the interface, the 48 trials per participant that were run were frustrating and difficult for some participants. If this number were increased, even using the proposed slider input instead of the binary search, a modest increase in number of data points (one or two) per best-fit linear function would be achieved. However, even a small gain would increase the confidence of the results and will be considered in the future.

Another consideration for sample size is the small number of participants. Only six participants performed this study. As such, this can only be considered a 'reality check', allowing the research to press onward, rather than a full confirmation.

6.4.4 Summary

In summary, enough evidence has been presented for this research to continue under the assumption that linear functions will work well for modeling colour differentiability. The amount of error that may be introduced is acceptable, and partially offset by the need to keep the calibration procedure short. Additionally, by having some measure of error introduced by this method, a confident prediction about the accuracy of the final predictor can be put forward. In the next chapter, the accuracy of the predictor is assessed to see how much credence can be placed in this accuracy prediction.

Chapter 7

Evaluation of Predictor Performance

In Chapter 5, the implementation of the model of colour differentiability was described. The basis for this model was the assumption of linear relationships, an assumption that was supported by the study in Chapter 6. The product of this model was a predictor that was hypothesized to correctly predict the point at which two colours will be considered distinct. The accuracy of this predictor is evaluated in the study presented in this chapter.

7.1 Goals

In this research, the differentiability of two colours is based on individual channel differences between the two colours. Given a colour, the predictor calculates how much each channel needs to be changed in order for the resulting colour to be considered different from the original colour. To evaluate the performance of this predictor, the following question is posed:

> How accurate is the predictor at predicting the point of differentiability for each channel of a given colour?

This study was designed to compare the differentiability limits selected by human participants against the limits generated by a predictor calibrated to each participant's colour differentiation abilities. These comparisons were performed on a randomly-generated set of colours from within the RGB colour cube. These random colours were generated once, and used repeatedly for each participant. The results of this study will provide a measure of the accuracy of the predictor proposed and developed in this research project.

7.2 Study Methods

7.2.1 Participants

Eight participants, all male, were recruited from the University of Saskatchewan. All were students or graduates from technical degree programs, and ranged in age from 19 - 29 years (mean age: 25 years). All had normal (or corrected-to-normal) visual acuity, and had no previous diagnosis or history of colour blindness. All participants were regular computer users (more than ten hours per week). Colour blind participants were not allowed to participate, as this study explored general accuracy of the predictor. Colour blind participants would potentially influence the results as the general model of colour differentiation (Chapter 4) and the calibration process for its implementation (Chapter 5) have not been specialized to handle the particular case of colour blind individuals. Further work is needed before colour blind participants can be modeled.

7.2.2 Apparatus

The study used a custom-built Java application. All participants performed the study using a laptop paired with an external CRT monitor running at a resolution of 1024 x 768. A CRT monitor was used because CRTs are less prone to angular colour variation - a frequent issue with LCD monitors, and a potential source of error in this study. The study was performed in full-screen mode on the external monitor. Participants interacted with the program using an external mouse. Keyboard input was not used.

7.2.3 Tasks

The software system described above was comprised of two separate stages. The first stage involved the calibration procedure outlined in Section 5.5.1 of Chapter 5. The end result of this calibration procedure was a predictor that provided a function that was evaluated in this study. This function accepts any RGB colour, then calculates and returns a set of six values. These values are the predicted upper and lower channel limits for the provided colour. The second stage of this study compared the predictions made by the predictor generated by the first stage to limits determined by human participants. The first stage will be referred to as the 'Calibration Stage', and the second stage will be referred to as the 'Evaluation Stage'.

In the evaluation stage, participants were presented with a rectangular field of circles, each approximately one `cm` in diameter. The task was divided into 72 individual trials. In each trial, the circles began as all one colour, but approximately half of the circles could have their colour modified using a slider at the top of the screen. Using this slider, the task was to identify the point at which the non-changeable colour and the slider-controlled colour became distinct from each other. When the participant identified this point, they selected a button labeled 'Proceed', ending the current trial, and another trial was presented. The participant was free to explore the entire range of possibilities in order to identify the point of differentiability. The interface for this stage is shown in Figure 7.1.

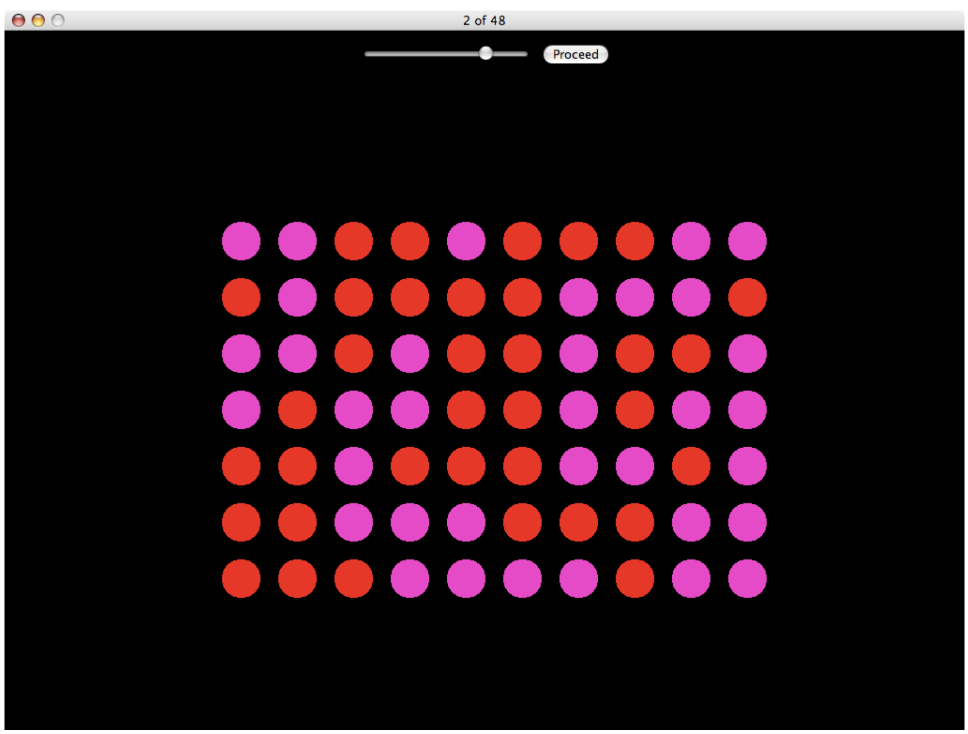

Figure 7.1: System used for evaluation stage of the study.

7.2.4 Procedure

Each study session involved a single participant at a time. The session procedure had three major steps; pre-study questionnaire, calibration stage, and evaluation stage. The calibration stage was composed of two separate components, each requiring participant involvement.

1. **Pre-Study Questionnaire**: This questionnaire was used to gather personal information about the participant. Personal information included age, gender, occupation, education level, computer experience (hours per week), electronic gaming experience (hours per week), information visualization experience (visualizations and tools used), corrected visual acuity (excellent, good, fair, poor), previous diagnosis of colour blindness, and family history of colour blindness. A sample pre-study questionnaire is given in Appendix A.

2. **First Calibration Stage**

   (a) **Delivery of Verbal Instructions**: The participant was informed that they could not fail at the task they were about to perform and that they could quit at any time. A short verbal description of the task was given as well.

   (b) **Program Instructions**: The first stage of the calibration process was started, and the participant was asked to read the written instructions for the study. The instructions for the first calibration stage are shown in Figure 7.2.

   (c) **Training**: The participant performed three trials of the first calibration stage to ensure they were confident in their understanding of the expectations from the task. The program was then stopped and restarted. No results were recorded from this step.

   (d) **Experimental Tasks**: The participant performed 12 trials for the first calibration stage.

(e) **Recording Results:** Once the participant completed the 12 trials, the results of this first calibration stage were recorded. If the participant had identified that any of the colour pairs presented in this stage were not different, the study was stopped, as this result suggested the participant had some form of colour blindness.

3. **Second Calibration Stage**:

   (a) **Delivery of Verbal Instructions**: The participant was informed that they could not fail at the task they were about to perform and that they could quit at any time. A short verbal description of the task was given as well.

   (b) **Program Instructions**: The second stage of the calibration process was started, and the participant was asked to read the written instructions for the study. The instructions for the second calibration stage are shown in Figure 7.3.

   (c) **Training**: The participant performed three trials of the second calibration stage to ensure they were confident in their understanding of the expectations from the task. The program was then stopped and restarted. No results were recorded from this step.

   (d) **Experimental Tasks**: The participant performed 48 trials for the second calibration stage, taking breaks every 12 trials to rest their eyes.

   (e) **Recording Results:** Once the participant completed the 48 trials, the results of this second calibration stage were recorded. These results were processed in a spreadsheet application to determine the inputs for the creation of the predictor. The predictor was constructed and given to the evaluation stage (described next).

4. **Evaluation Stage**:

   (a) **Delivery of Verbal Instructions**: The participant was informed that they could not fail at the task they were about to perform and that they could quit at any time. The participant was informed that the task for this stage was identical to the task for the second calibration stage.

(b) **Program Instructions**: At this point the evaluation stage of the program was started, and the participant was asked to read the written instructions for the study. These instructions are shown in Figure 7.4.

(c) **Training**: As the task for this stage was identical to the second calibration stage task, no training was needed.

(d) **Experimental Tasks**: The participant then performed 72 trials, taking breaks every 12 trials to rest their eyes.

(e) **Recording Results:** Once the participant completed the 72 trials, the results of this third stage were recorded for processing.

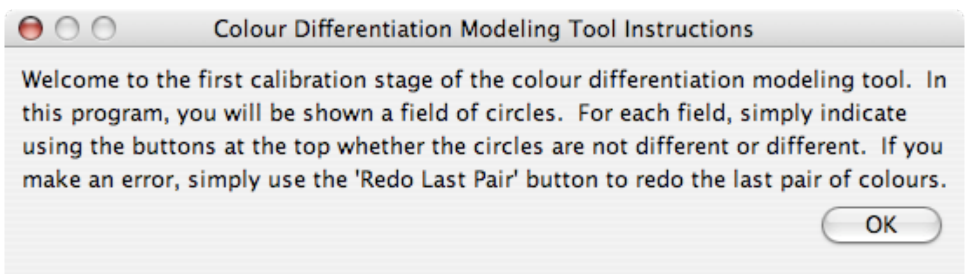

Figure 7.2: Instructions given at the beginning of the first calibration stage of the study.

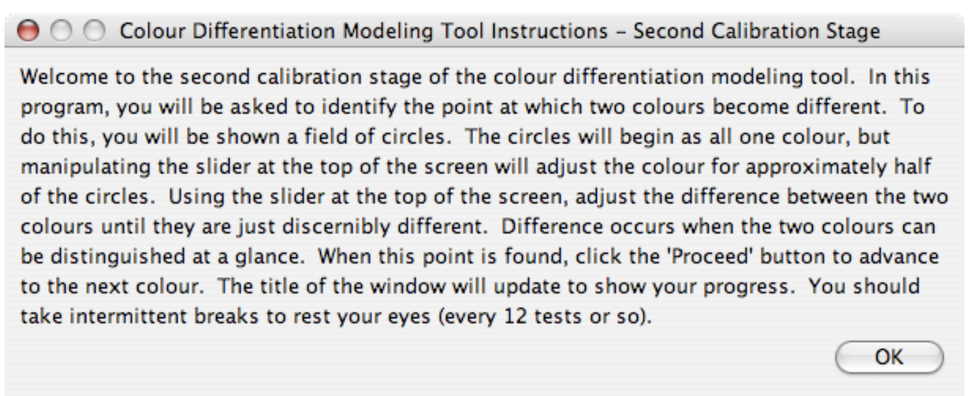

Figure 7.3: Instructions given at the beginning of the second calibration stage of the study.

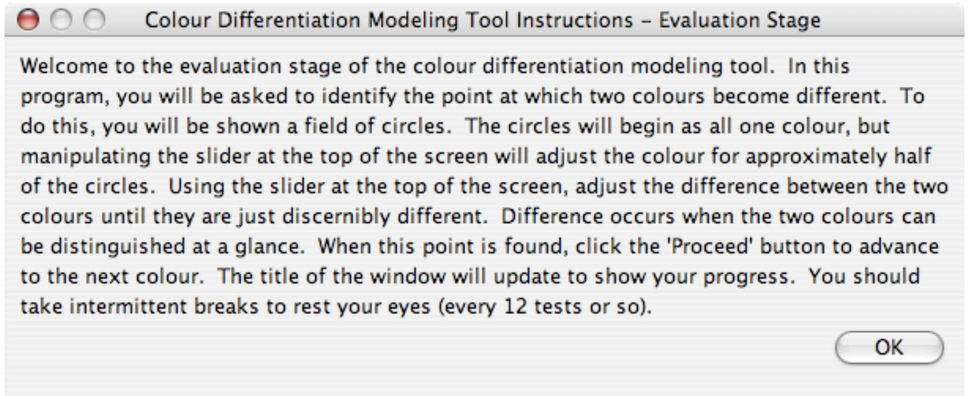

Figure 7.4: Instructions given at the beginning of the evaluation stage of the study.

7.2.5 Study Design

In this study, no colour blind participants were used. This study was designed to evaluate the effectiveness of the colour differentiation predictor. To simplify this assessment process, it was decided that the colour perception of the participant should be controlled as much as possible. Due to the potential variation in type and severity of colour blindness (discussed in Section 2.4), the simplest control would be to only allow non-colour blind individuals to participate in the study.

This study was split into two major stages: calibration and evaluation. The calibration stage provided necessary information for the evaluation stage (namely the predictor to be evaluated). In the following sections, the study design for each of these stages will be discussed.

Study Design - Calibration Stage

The details for calibration were given in Section 5.5. As an aid to the reader, a brief overview of the calibration procedure is given:

1. Assessed whether the participant could differentiate between the corners of the RGB colour cube. The corners compared were those that lie on each end of an exterior edge of the RGB colour cube.

2. Measured the differentiation limit for each channel of each colour represented by

the corners of the RGB colour cube. To increase the confidence in the measures recorded, each corner measurement was repeated twice and the resulting two values were averaged.

3. When a channel value was at its maximum value (255), it was not possible to increase the channel value from this point to obtain an upper limit. Instead, a lower limit was measured in step 2 and extrapolated within a spreadsheet program to obtain an upper limit.

4. The values obtained from the extrapolation step were then used to construct a predictor. A predictor contained the definition of every function that defined differentiation along an outside edge of the RGB colour cube.

Study Design - Evaluation Stage

The purpose of the evaluation stage was to compare the limits generated by the predictor to limits gathered directly from the participant. To compare these limits, a set of 12 colours were used. Each channel for each of these colours was randomly selected from the range [0-127]. The reason this range was restricted to [0-127] instead of [0-255] is explained below. These 12 colours are listed in Table 7.1. As the results for each participant were to be aggregated for analysis, every participant performed the evaluation stage using the same set of 12 base colours, although the order in which colours were presented was randomized for each participant. The results from this stage are used to answer the question posed in Section 7.1.

Base Colours (R,G,B)			
(102,82,103)	(0,103,112)	(61,98,119)	(55,109,0)
(100,110,0)	(118,0,42)	(0,125,41)	(42,16,33)
(35,0,85)	(64,48,78)	(77,52,70)	(119,69,77)

Table 7.1: Each of the randomly-selected base colours for the evaluation stage of the study.

Each of these 12 colours served as the base colour in a trial. Limits were gathered for each channel of each of these colours to give a total of 36 trials. Each trial was repeated twice to give an indication of participant accuracy. Participant accuracy results are presented in Section 7.3.2. This resulted in the following number of trials:

$$
\begin{array}{rl}
 & 12 \text{ base colours} \\
\times & 3 \text{ channels per colour} \\
\times & 2 \text{ repetitions of each channel} \\
\hline
 & 72 \text{ total trials}
\end{array}
$$

For a given trial, the participant was presented with the interface shown in Figure 7.1. The base colour came from the set of 12 random colours, and a slider manipulated the value of one of the channels. At the extreme left, the slider contributed nothing to the channel (and the colours were identical). As the slider was moved to the right, an increasing amount was added to the channel the slider was tied to. At the extreme right, the value for the channel was at 255, representing the maximum possible difference between the base colour and the slider-controlled colour. The participant explored the potential colours until the point of differentiation was found. Then the participant selected the 'Proceed' button, and advanced to the next trial.

For each comparison between participant limits and predictor limits, only upper limits were used. This allowed more colours to be used (as this simplification halved the number of potential trials). Using more colours provides greater confidence in the results of the comparisons. In Section 4.5, it was described how lower limits can be deduced from upper limit values. Because of this it was deemed unnecessary to test lower limits.

As only upper limits were compared, the non-changeable colours needed to have their channel values kept artificially low. In this study, it was decided that the range [0-127] provided enough 'headroom' for participants to find the upper limit for any given channel. As upper limits involve increasing a channel value until differentiation is achieved, capping channel values for the non-changeable colours to 127 worked well and no participant chose a maximum possible upper limit. A maximum possible upper limit could indicate that a sufficient difference was not achievable.

For a given trial, several values were recorded. The channel values for the non-changeable colour served as the base value for each trial, so these were recorded. The 'channel of interest' is the channel which is controlled by the slider. This was recorded. The non-changeable colour was given to the predictor, and the predicted upper limit for the 'channel of interest' was recorded. Finally, the participant-selected upper limit for the 'channel of interest' was also recorded. These values allow a comparison between the predicted and actual values to be performed. This comparison provides a measure of the accuracy of the predictor.

7.3 Results

In Section 7.1 the following question was posed:

> How accurate is the predictor at predicting the point of differentiability for each channel of a given colour?

In this section, the general results of this study are presented. After this, intra-participant variation is explored.

7.3.1 General Results

In this study, eight participants performed 72 trials each, giving 576 individual measurements recorded. For each of these 576 trials, the predicted upper limit (the value generated by the predictor) and the participant upper limit (the value provided by the participant) were recorded, and are compared here. To compare the two upper limits for each trial, the predicted upper limit was subtracted from the participant upper limit to get the difference between them. When this value is negative, it indicates that the predictor predicted a limit that was too high. When this difference is positive, the predictor predicted a limit that was too low. In Table 7.2, some high-level properties of these error values are presented.

It is interesting to note that the standard deviations are almost the same for each channel, but the means and medians vary considerably. For the red and green

Channel	Minimum	Maximum	Median	Mean	Standard Deviation
Red	-45	36	-10	-9.5	14.4
Green	-66	25	-12	-13.3	13.8
Blue	-25	52	1	2.7	13.6
Overall	-66	52	-7	-6.7	15.5

Table 7.2: Summary statistics for the prediction error between the predicted limits and participant limits. Red, Green, and Blue indicate the aggregate difference summary statistics for all trials performed on these channels. Overall indicates the summary statistics for the entire data set.

channels, the predictor seems to be consistently over-predicting. This is illustrated not just by the means (-9.5 and -13.3, respectively), but also by their respective ranges between minimum and maximum values (-45 → 36 for red, and -66 → 25 for green). The blue channel seems biased in the other direction (under-predicting). The mean for this channel is 2.7, with the range between minimum and maximum values shifted the other way (-25 → 52).

The results presented in Table 7.2 suggest that the errors are normally distributed (mean of -6.7 and standard deviation of 15.5). The distributions for each of the data sets presented in this table are shown in Figures 7.5 - 7.8. As can be seen, each distribution seems to be normal, although the green and blue channel differences are heavier on the negative side than the positive. This bias affects the overall differences as well.

7.3.2 Intra-Participant Variation

As mentioned above (Section 7.2.5), it was desired to see how consistent a participant was between repetitions of a particular trial. To measure this consistency, the participant limit values for each repetition were compared. As there were exactly two repetitions of each (non-changeable colour + channel) pair, this gave two upper limits for exactly the same situation.

To get a measure of intra-participant consistency, these two values were sub-

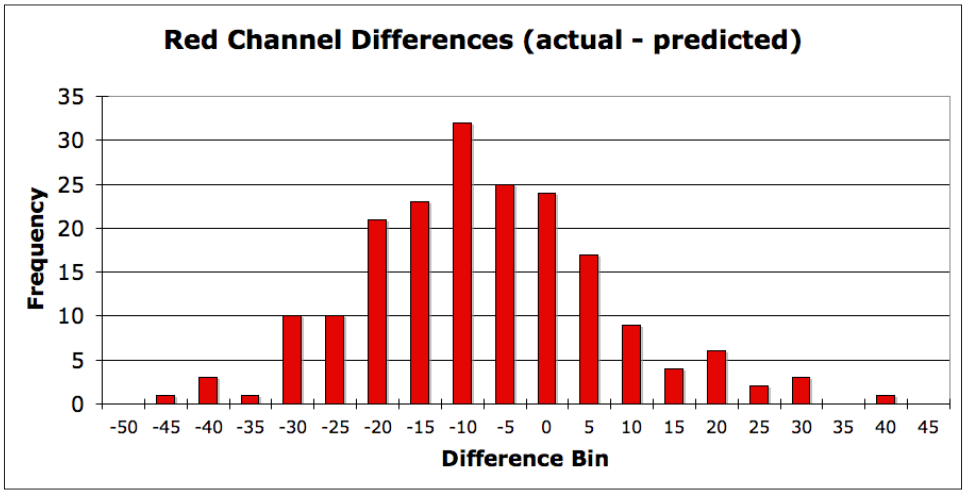

Figure 7.5: Red channel differences between the predicted upper limit and the participant upper limit.

tracted from each other. As the order of subtracting these two limits is irrelevant (it would simply swap the signs for all the differences), the second limit was subtracted from the first limit. Some overall characteristics of these subtracted values are presented in Table 7.3.

The distribution for intra-participant variation is given in Figure 7.9. It can be seen that the resulting distribution is centered near zero (median = -2 and mean = -2.8), and the distribution drops off rapidly from this central point. There are some extreme values (and these account for the extreme values for minimum and maximum values in Table 7.3.) As there are 576 trials in total, this gives 288 pairs of values that are for the same trial data (repetitions). Of these 288 pairs, 203 of them lie in the range [-10, +10].

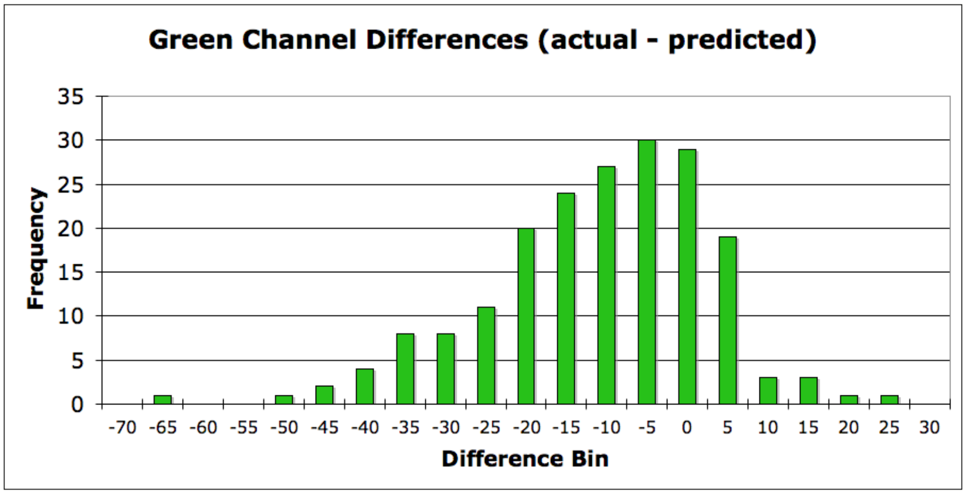

Figure 7.6: Green channel differences between the predicted upper limit and the participant upper limit.

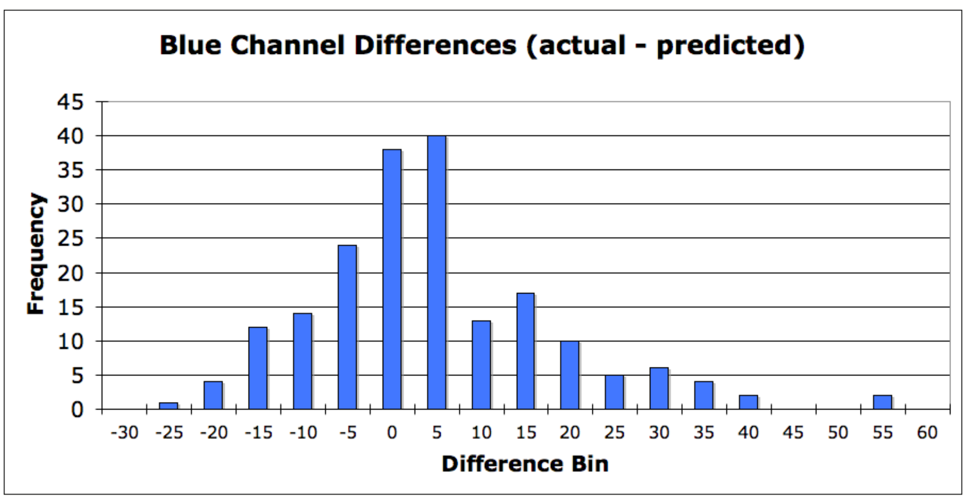

Figure 7.7: Blue channel differences between the predicted upper limit and the participant upper limit.

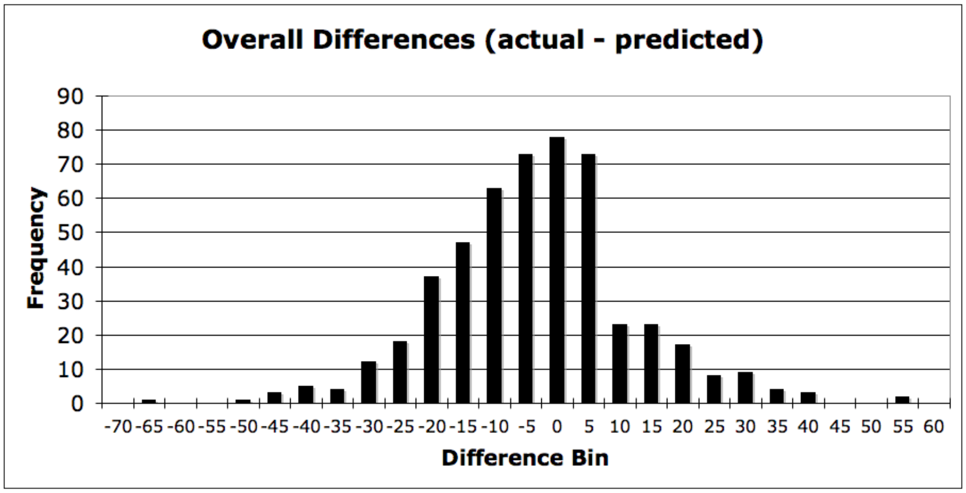

Figure 7.8: Overall differences between the predicted upper limit and the participant upper limit.

Participant	Minimum	Maximum	Median	Mean	Standard Deviation
1	-32	40	0	-0.7	13.4
2	-8	11	0	0.6	4.0
3	-44	13	-5	-10.7	14.2
4	-24	24	-1	-2.0	8.1
5	-17	14	-0.5	0.0	7.3
6	-42	43	0	-1.1	18.1
7	-28	19	-6	-6.3	10.5
8	-18	15	-2	-2.4	5.7
Overall	-44	43	-2	-2.8	11.5

Table 7.3: Summary statistics for intra-participant consistency testing. These are the minimum, maximum, median, mean and standard deviations per participant, as well as overall.

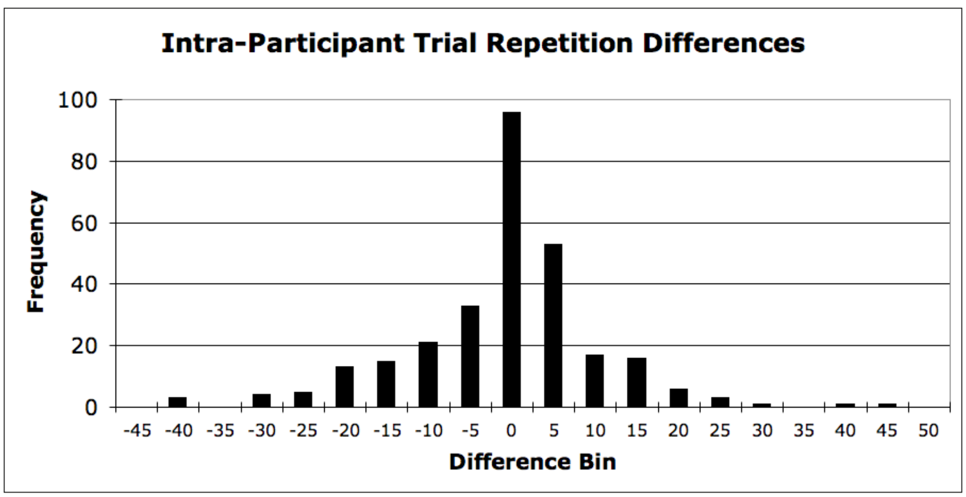

Figure 7.9: Overall intra-participant variation. The difference for each pair of limit values (two repetitions) was calculated and the results were charted in this histogram.

Chapter 8
Discussion

In this chapter, the results are analyzed to determine what scientific contribution has been gained. To do this, the results of the research will first be reviewed, and then the principal problem posed in Chapter 1 will be considered using these results. The process used to conduct this research will then be reviewed in terms of factors that may have influenced the results. The chapter ends with a discussion regarding the generalization of the research and a review of design decisions made throughout this research.

8.1 Summary of Results

In Chapter 6, a study was performed to evaluate the usefulness of a linear function for predicting differentiation limits for colours. This study found that linear functions would work well, subject to a range of confidence. This range was found to be [-14.7, 14.7], and means that when a linear function is used to predict a differentiable limit, 95% of the time the actual limit will be within $\pm$ 14.7 intensity values of this prediction. This essentially means that the predictor will be right 95% of the time, as long as 'right' is defined such that the actual value falls within this range.

In Chapter 7, another study was performed to evaluate the accuracy of the predictor described in Chapters 4 and 5. The results from this study are that on average, the predictor over-predicted (mean difference (actual-predicted) = -6.7), and had an overall standard deviation (SD) of 15.5. This gives a 95% (two standard deviation) range of [-37.7, 24.3]. This means that 95% of the predictions made by the predictor fell within the range [-37.7, 24.3] from the actual, meaning that in 95% of the cases, the predictor over-predicted by a maximum of 37.7, or under-predicted by a maximum of 24.3.

As a secondary element in Chapter 7, an examination of intra-participant variation was performed. As each unique trial was repeated twice for every participant, an indication of how consistent a participant was between these two repetitions could be gained. Overall, the average difference was -2.3, with a standard deviation of 11.5. The minimum difference was -44, and the maximum difference was 43. This means that in general, the participants were quite consistent, but there were cases where the participant was not very consistent at all (judging by the minimum and maximum values). In total, there were 288 of these differences calculated, and 203 of them were within ten values of each other. Overall, $\sim 70\%$ of the participant responses to two identical trials were within ten values of each other.

8.2 Implications of Results

The implications of these results are now explored. This is done by examining the original problem posed in Chapter 1 and exploring how these results address this problem

8.2.1 Original Problem

The problem presented at the beginning of this thesis was:

> Current colour perception models fail to accurately predict the colour confusion problems for a particular individual because they are not sensitive to the specific context of a colour perception task.

As a solution to this problem, this thesis proposed the construction of a model of colour differentiation which accurately represents how a particular individual differentiates between colours in a particular colour perception environment. This model was constructed in the context of categorical encoding.

Ultimately, a summary of this thesis is: Colour is used to represent information. Often, colour is used to encode a *category*, such that all the elements of a single colour belong to the same category (categorical encoding). This use of colour relies on individuals begin able to tell the difference between two colours (differentiation). Can a model that represents how well an individual differentiates between colours be constructed? If it can, then the model can be used to predict when two colours are *not* going to be differentiated by the individual. When this situation arises, the model can also be used to find *other* colours that will be differentiable.

8.2.2 Accuracy of the Predictor

The evaluations show that a model that represents how well an individual differentiated between colours can be constructed. The accuracy of this model is now explored.

In Chapter 7, this predictor was evaluated by comparing the limits it generated to limits selected by the participant for a set of random colours. These were referred to as predictor limits and participant limits. The difference between the predictor limits and the participant limits was recorded. The results of all of these comparisons were analyzed to see how much the differences varied. The means and standard deviation for these differences are given in Section 8.1.

As can be seen in Figure 8.1 (replica of Figure 7.8), the differences between the participant values and predictor values (participant-predictor) follow a somewhat normal distribution, with a mean of -6.7, and a standard deviation of 15.5.

To determine the accuracy of the predictor, a range of values can be established that defines the acceptable error rate. This range would specify the minimum and maximum allowed error, such that all values between these extremes are considered accurate. The error has been calculated as the participant-predictor difference. Using this range, all participant-predictor difference values can then be labeled in a binary fashion: either they are considered accurate or they are not. When a difference is considered accurate, this means that the predicted value is close enough to the participant value to be considered accurate. When a difference is not considered accurate, this means that the predicted value is too different from the participant value, and therefore not considered accurate. With this binary labeling, an overall

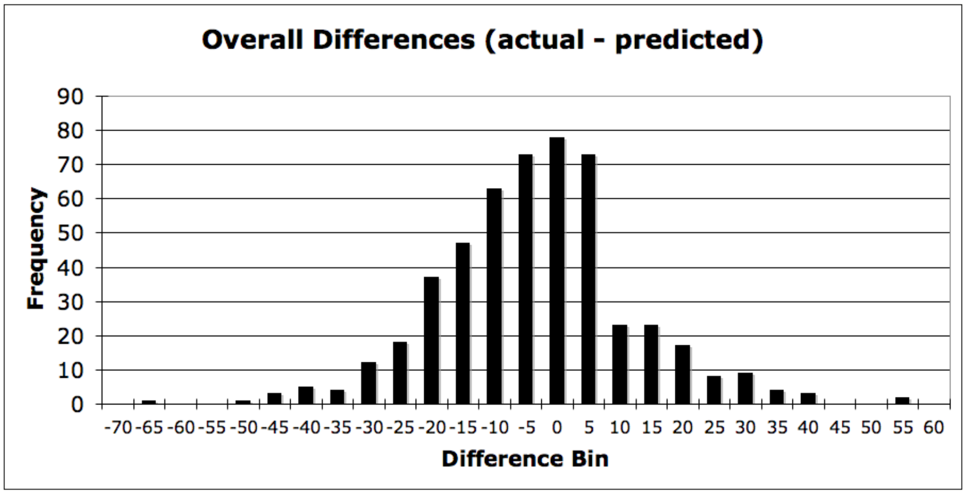

Figure 8.1: Overall differences between the predicted upper limit and the participant upper limit (participant-prediction).

percentage accuracy for the predictor can be constructed by dividing the number of accurate predictions by the total number of predictions.

One way to define this accuracy range is to use the intra-participant variation data from Chapter 7. The intra-participant variation data illustrates the potential variation of an individual when identifying the differentiation limit for repeated experimental trials. This data follows a normal distribution as shown in Figure 8.2. Using the mean and standard deviations for this distribution, the range which encompasses 95% of all intra-participant variation can be calculated. 95% of all intra-participant variation will be covered by the range [mean - (two $\times$ standard deviation), mean + (two $\times$ standard deviation)]. These values are -2.3 $\pm$ (two $\times$ 11.5) = [-25.4, +20.8]. 95% of all intra-participant variation values fall within this range, this can be used to evaluate the 'close enough' accuracy of the predictor. This can be used because the predictor *cannot* be expected to be more precise than the individual - i.e., if the individual varies by some amount between identifying limits for identical situations, then it is reasonable to assume that the predictor is close enough if it falls within this variation.

To do this, a frequency count of all participant-predictor values that fall within the range [-25.4, +20.8] is conducted. As the participant-predictor values are whole numbers, the actual range is [-25, 20], but this small change should make little differ-

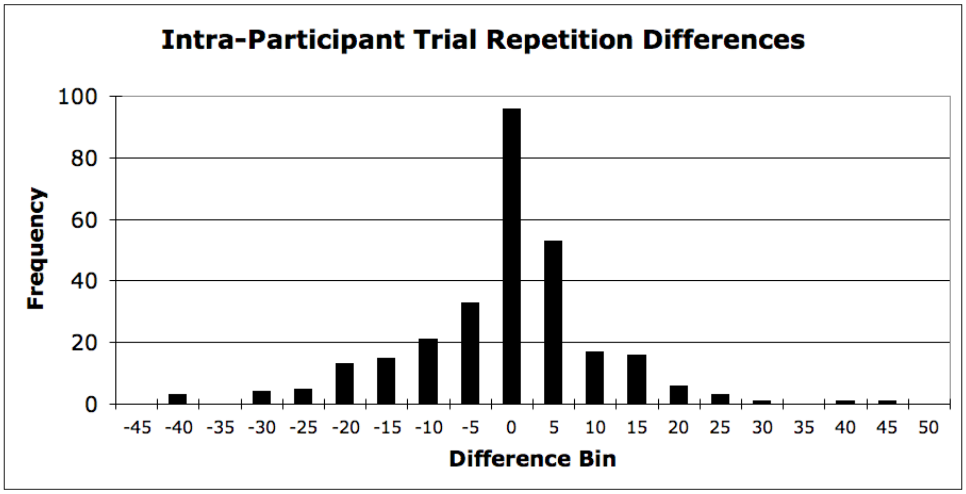

Figure 8.2: Overall intra-participant variation. The difference for each pair of limit values (two repetitions) was calculated and the results were charted in this histogram.

ence. There were a total of 576 differences gathered during the study in Chapter 7, and of these 497 fall within the range [-25, 20]. This means that by this metric, 497 predictions made by the predictor are correct. This gives an accuracy of 497/576 = 86.3%.

Based on the findings of intra-participant variation, the predictor proposed in this thesis makes a differentiability limit prediction that is as accurate as people themselves 86.3% of the time.

8.2.3 Implications for Original Problem

This thesis demonstrates that an accurate model that represents how well an individual differentiates between colours can be constructed. How does this success reflect back on the original problem (shown below)?

> Current colour perception models fail to accurately predict the colour confusion problems for a particular individual because they are not sensitive to the specific context of a colour perception task.

Current models of colour differentiation fail because they do not accurately model colour perception in terms of the context of the colour perception. The current models of colour differentiation are not sufficiently sensitive to the factors that influence colour perception (Chapter 2).

This thesis describes a model that *does* account for all of the potential factors that influence colour perception. Accounting for these factors is accomplished through the use of a *judgement* task to construct the model. An individual supplies calibration information for the model through a series of colour tasks. Using this calibration information, a predictor (encoding the model) can be created. As shown above, this predictor achieves an accuracy of 86.3%.

Accuracy of the traditional model of colour differentiation is unknown. It is known that this model relies upon certain assumptions in order to operate accurately. These include the use of a calibrated monitor, controlled lighting situations, and full knowledge of the type and severity of any colour blindness present. It is also assumed that any and all of the factors that influence colour perception (Chapter 2), are either controlled or known. It seems reasonable that this model will work well in the situations where these assumptions are indeed correct. However, situations where these assumptions are incorrect may cause this model to not work well. How well this model works could be defined in terms of accuracy as given here (i.e., how many times the model makes a correct prediction of differentiability divided by total predictions).

The contribution of the model described in this thesis is that applications that currently rely on the traditional model of colour differentiation (see Section 2.8) now have another model to base their colour adaptation strategies upon. The presence of two models of colour differentiation allows three possible strategies:

1. The model presented in this thesis is superior in comparison to the traditional model, and should be used exclusively by colour adaptation applications.

2. The traditional model is superior in comparison to the model presented in this thesis, and should continue to be used.

3. Each model is superior to the other under certain circumstances. In this case, colour adaptation applications can take a hybrid approach, using the traditional model when conditions favour it, and using the model presented in this

thesis when conditions favour it.

Both cases #1 and #2 are unlikely. The model presented in this thesis should outperform the traditional model when factors are unknown or uncontrolled (as in typical computer use situations). As the accuracy of the mathematical model is unknown, it may exceed the accuracy of the model presented in this thesis when all of its assumptions are correct, although this situation would be very rare in real-world computing situations. As a result of this possibility, Case #3 seems to be the most likely to produce the best adaptation results.

A hybrid approach would allow a broader set of circumstances to be handled by the colour adaptation tools. When traditional model's assumptions hold, it could be used. In situations where its assumptions do not hold, the model presented in this thesis could be used. The end result is that colour adaptation applications will benefit from a colour differentiation modeling strategy that handles a broader set of colour perception circumstances. This will ultimately provide better tools for assisting individuals who experience atypical colour perception when using colour-encoded information visualizations. In terms of the original problem, this is the ultimate contribution of this research:

> The colour differentiation model presented in this thesis can be used to accurately predict the colour confusion problems for a particular individual because it is sensitive to the specific context of a colour differentiation task.

8.3 Limitations and Critical Reflection

In this section, limitations and strengths of this thesis will be explored. Critical reflection on design decisions made during this project are also presented.

8.3.1 Limitations

Some limitations of the system presented in this thesis are explored now. These limitations are the potential for frequent calibrations, and the work involved in expanding this model to colour use techniques beyond categorical encoding.

Frequency of Calibrations

The system presented in this thesis requires the use of a calibration step in order to construct the predictor. The use of judgement tasks for this calibration allows the automatic consideration of all of the factors that influence colour perception and differentiation. As a result, the calibrated model is tuned to a particular context (environment and person). When this context changes, the model may become invalidated. As an example, consider the model constructed for a particular monitor. If the individual for whom the model was created moved the model to another machine with a different monitor, the context has changed and therefore the model may be invalidated, causing it to lose accuracy. Throughout the development of the calibration system, calibrations were performed on a variety of monitors (Dell 17" external CRT monitor, 15" Apple MacBook Pro notebook screen, 19" Samsung and 22" Dell external LCD monitors), and it was found that the calibration data changed substantially from monitor to monitor.

When the model was evaluated in Chapter 7, care was taken to perform the evaluation stage *immediately* after the calibration stage, while maintaining the status of the environment as much as possible. To do this, no lights were turned on or off during the process, and blinds were kept open or closed throughout. Performing the evaluation immediately after the calibration reduced the chance of any other changes such as ambient lighting levels.

No evaluations have been performed to determine the accuracy of the predictor when the calibration and the evaluation are performed in different situations. Because of this, not much can be said about the generalizability of the findings (i.e., will one calibration work for all situations?). To test this, a calibration would need to be performed in one context, and then evaluated both in that context (to form a baseline), and again in another context, or a series of different contexts (to allow comparison).

As a result of this missing research, it is unknown how frequently an individual would need to re-calibrate the model. A potential solution could be that a battery of calibrations could be performed that cover a wide range of contexts. This battery of calibrations would only need to be performed once. In the future, when a new context arises, the calibration that was performed in the context that most closely matches the new context could be used to replace the existing calibration.

Lack of this knowledge imposes a limitation on the current system. Calibrations are time-consuming, taking approximately 20-30 minutes, and can be tiring. If calibrations need to be performed frequently, the benefit provided by them may be outweighed by the losses incurred by the frequent calibrations.

Application to Additional Colour Tasks

In this thesis, the particular colour task of categorical encoding in a scatterplot-like visualization is examined in detail. It seems safe to presume that the approach taken here would work equally well for other colour tasks in which colour is used as a label (highlight, popout, brushing - see Chapter 3), as each of these relies on differentiability between colours. It is unknown how the model presented in this thesis would extend to other tasks, such as colour as value (false colour, continuums, and multi-dimensional data display), and colour as imitation of reality. Colour as value uses colour differentiation as well, so it should be able to take advantage of the modeling approach presented in this thesis. Additional research is necessary before extending the findings of this thesis to additional colour use techniques in information visualization.

8.3.2 Strengths

Some strengths of the system presented in this thesis are explored now. These strengths include the ease of integrating this modeling approach into existing applications that perform colour adaptation, and the superiority of this model over the traditional colour modeling approach.

Integration into Existing Colour Adaptation Tools

As existing colour adaptation tools employ a simple interface when interacting with the currently used traditional model of colour perception, it seems reasonable that they could be extended to use the model presented in this thesis.

Currently, the predictor in this thesis provides two colour differentiation functions. The first is what is used to evaluate the predictor in Chapter 7. This function accepts a colour specified as a RGB triple of values between 0 and 255, and returns a sextet of values, representing the three upper and three lower limits for that given colour, an upper and lower limit for each channel. The second function is an extension of the first function that provides a more sophisticated interface for adaptation tool developers. In this second function, two colours are given to the predictor (expressed in the same way as above), and a boolean response is given back. This boolean response simply indicates whether the two colours are differentiable, according to the model of colour differentiation encoded in the predictor.

These functions could be extended further to provide an even more useful set of services. These extensions include a function that takes any number of colours, and provides another colour that is differentiable from the provided colours. To make this even more useful, the colour task being performed could be given to the function as well. The colour task may be 'categorical' or 'popout' and a suitable colour found according to this parameter.

Another extension would be to provide a function that accepts a single integer argument, and returns a set containing the specified number of colours. These colours will be selected such that they are maximally differentiable from each other. Of course, this task becomes more difficult as the input parameter increases, so some reasonable limit would need to be imposed (or the user may have to accept that there may be indifferentiable colours in the set).

All of these extensions would not be too difficult to implement, and would provide greater power to the developers of colour adaptation applications. Using some of the functions described above, the model could even provide entire colour schemes for particular visualization techniques.

Superiority of This Model in Particular Contexts

A main strength of this system is in situations where the context for colour differentiation falls outside of the assumptions required by the traditional model of colour perception currently used. Using a judgement-based calibration procedure, the model presented in this thesis can be tuned to any situation. Although coverage testing has yet to be performed for a broad range of situations, it is hoped that the accuracy score of 86.3% determined in this thesis should be consistently achieved. As such, the system presented in this thesis should be accurate in many contexts, subject to repeated calibration, as described above (Section 8.3.1).

8.3.3 Critical Reflection

In this section, particular design decisions made during this research will be given a brief explanation and examined. This includes some discussion of what could have been done differently, including reconsideration of the assumption of linearity, the use of a black background in calibration and evaluations, properties of the field of circles, and the use of a judgement versus a performance task.

Linear Assumption

To model colour differentiation, a set of linear functions was used. The linear assumption and its benefits outlined in Section 5.1 greatly simplified the calibration process, which allowed calibration to be performed in a reasonable amount of time. The assumption of linearity was also supported by the results of the study described in Chapter 6. It is possible, however, that linear functions may not be the best functions to describe colour differentiation.

Perhaps a more complex function could be used to describe colour differentiation more accurately. Candidate functions would include quadratic functions and piecewise linear functions. Quadratic functions provide a more expressive language for describing relationships in the form $f(x) = ax^2 + bx + c$. More research would be required to determine the minimum number of points necessary to define a quadratic function (if there is one). Even though these would be potentially more expressive, quadratic functions may add an excessive number of measurements to the calibration stage.

Another possibility would be to gather midpoints along each outer edge of the RGB colour cube. These values could be used to generate piecewise linear functions along each edge. In cases where the edge *is* defined by a single linear function, piecewise functions will introduce no additional difficulty. In cases where the relationship is not linear, piecewise linear functions would allow a more accurate description of the relationship. This has an added benefit of allowing a tunable number of calibration measurements to be made. Every increase of a midpoint value would add 12 measurements to the calibration process, so the person performing the calibration could specify how many measurements should be made (corresponding directly to how long the calibration process will take).

A third option is supported by adding more points in a controlled fashion as described above for piece-wise linear functions. Instead of constructing piece-wise functions, the additional points could be used to better inform a least-squares method of finding the linear function that best fits the datapoints. This approach would improve the accuracy of the linear function approach, but still allow simplified interpolation procedures as outlined in Chapter 5.

Black Background

The calibration and evaluation software systems all used a field of circles set against a black background. It was hypothesized by one of the participants that the black background could be biasing the results of the study - the participant felt that if the background was white, different results could be found.

The reason black was chosen as a background colour was due to the additive nature of colour in digital environments. As black is the colour produced by the lack of stimulation of monitor pixels, it was decided early that this would be the optimal background colour, as it would limit unwanted stimulation of participant cones. Although a consistent background colour would not influence the accuracy of the predictor (as it would not change from calibration to evaluation), it is possible that the background colour would influence the shape of the linear functions used for prediction.

Although the decision to use a black background still seems to be the correct choice, there remains the possibility that it adds another factor to the prediction model. The predictor assumes a black background for all differentiation tasks, therefore the background colour is not a parameter for any of the functions supplied by the predictor. As a minimum, it may be necessary to specify the background colour in any calls made to the predictor.

Field of Circles

Many potential complications reside in the use of the field of circles to measure colour differentiation. Throughout this research, many different configurations were used in attempts to make this measurement. It was found that differentiation abilities are greatly enhanced by proximity, such that if the two colours to be discerned are adjacent (with no background colour between them) it is much easier to detect differences.

The use of circles was originally inspired by the circular shape of the fovea of the eye (the highest resolution and most colour-sensitive region of the retina). The fovea contributes a very small amount of the total visual range (about $2°$ of the total field of view), and it was decided that a circular patch of colour would most effectively utilize the fovea - allowing the greatest assessment of colour. Originally, one large circular region was used. This region was divided into two equal-sized semicircles. Between the semicircles was a strip of background colour. The colour of one semicircle was participant controlled (in a manner similar to the slider used in Chapter 7), and the other semicircle was the base colour. Although this setup worked well initially, it quickly resulted in difficulties. These difficulties arose from fatigue of the cones of the eye, causing an afterimage effect in which the changeable colour appeared darker than the non-changeable colour when they were actually the same colour. This naturally affected the results of this preliminary work, so another solution was sought out.

The solution eventually decided upon was the field of circles illustrated in Chapters 5, 6, and 7. Although this reduced the severity of afterimage difficulties, they were still present if the same pattern of colours was presented to the user. To help with this, randomization was employed. Approximately half of the circles were one colour and the remainder were another, after each manipulation of the colours by the participant, these approximate halves were reassigned. Essentially for each manipulation, as a circle was drawn, it was decided using a boolean random value whether the circle should be one colour or the other. Using this technique, the two colours transitioned through random 'patterns' as the participant manipulated the colour. This considerably reduced any remaining afterimage difficulties, and getting the participant to take a break every 12 trials further reduced them.

Another consideration is the distance between the circles. As described above, circles that are not separated by any background space allowed participants to find minute variations in colour (probably due to contrasting intensity). This led to the introduction of a gap between circles. The size of this gap was bounded by the foveal area of focus described above, and the desire to eliminate the evaluation of colour memory. Essentially, the circles had to be close enough to allow quick scanning between circles of different colours (within the foveal regions). The circles could not be so far apart as to require the individual to 'remember' the colour when scanning from circle to circle.

Ultimately, the research results may have been influenced by the choice in gap size. Further work is needed to identify this influence.

Judgement Versus Performance

The task the participants performed during calibration and the evaluations was a *judgement* task not a performance task. Selecting the point at which two colours became distinct through the use of a slider requires some form of judgement on behalf of the participant. This is because the line between not differentiable and differentiable is not clear cut, but transitions through a ‘fuzzy’ zone. This is confirmed by the intra-participant variation explored above (Section 8.1). As a result, the participant had to determine when they thought the colours were distinctly different through a judgement process.

Using a performance task would require a different approach. Instead of using a slider to allow the participant to explore the available colours, a performance approach would present the participant with a series of yes/no questions - similar to the first stage of the calibration process (Section 5.5.2). Essentially, a series of circle fields of varying colours could be shown to the participant. For each field, the participant would simply indicate whether the two colours are different or not different. This approach has an inherent difficulty, however. Because of the ‘fuzzy’ zone between *not different* and *different*, there will be a set of circles which will be somewhere in between these two labels. For identical circle fields from this ‘fuzzy’ zone, the response from the participant will one time be ‘different’, and another time be ‘not different’. As a result, multiple trials will be needed to establish the true boundaries of this ‘fuzzy’ zone, as well as the nature of this zone. In addition to the added strain of performing multiple trials, some decision will need to be made as to where in this ‘fuzzy’ zone should the limit be defined.

As an example, consider a scenario in which the participant is determining an upper limit along the red channel. A series of test circle fields are generated in which the channel value increases by a value of one for each field. These fields are randomized and shown to the participant. Their response for each field is needed. To help identify the ‘fuzzy’ range, perhaps five repetitions of each field are performed. Imagine that the red channel for the non-changing colour is at 154. This would result in (255-154) = 101 different fields $\times$ five repetitions = 505 separate trials - all for only a single upper limit measurement! On top of this, there is the difficulty of determining where the limit is. Suppose a ‘not different’ is recorded as having a value of zero, and a ‘different’ is recorded with the value one. For each channel value, the participant responses are summed and a selection of these are presented in Table 8.1. Where should the limit value be for this data set?

164	165	166	167	168	169	170	171	172	173	174	175	176	177
0	0	1	2	0	2	4	2	3	5	5	4	5	5

Table 8.1: Hypothetical results for a pure performance evaluation. The value for the red channel (the channel of interest), is on top, and the sum of the participant responses is on the bottom.

The judgement approach solves both the tedium of numerous trials, and the uncertainty about the location of the limit. This is accomplished by allowing the participant to explore all of the potential colours, and letting them *choose* where the limit point is. All that is required is that the participant identifies that there is some room for judgement here, and that they try to remain consistent from trial to trial. If five repetitions of the colour field are performed, the number of trials is reduced from 505 to 5! To determine the limit, the mean, median, or mode could be calculated for the resulting five values. This is why the judgement approach was chosen for the implementation.

Chapter 9

Conclusion and Future Work

9.1 Summary

There are many factors that influence the colour perception abilities of humans, leading to atypical colour perception. This can cause problems when colour is used to encode information in visualizations. When an individual is experiencing atypical colour perception, the use of colour encodings can be perceived incorrectly, leading to mistakes and frustration. Colour adaptation tools attempt to automatically recolour a visualization to alleviate this problem, but these tools currently rely upon a rigid model of colour perception that rests on several assumptions; assumptions that are often not met in common colour-viewing environments. The failure to meet these assumptions causes colour adaptation tools to be less effective.

In this thesis, a solution to this limitation of current colour models was proposed. This solution provides a colour differentiation model that is based on human judgement in the setting of interest, not on assumptions. As a result, the model is tuned to a particular colour viewing environment. This tuning affords precision for the proposed model, which was 86.3% as accurate as human judgement of differentiability in the studies presented in this thesis.

Overall, predicting colour differentiation using a model that automatically incorporates contextual factors was successful. This success should allow colour adaptation tools to perform better in more colour perception contexts than currently possible. Improving the performance of colour adaptation tools should decrease mistakes and frustration for individuals experiencing atypical colour perception as they use colour-encoded information visualizations.

9.2 Contributions

The main contribution of this thesis is the idea of contextualizing the model of colour differentiation and the development of this model. Secondary contributions include an evaluation of the assumption of linearity, an evaluation of the model-based predictor, a taxonomy of colour use in information visualization, an analysis of internal and external factors that influence colour perception, and software implementations of the calibration procedure, the model-based predictor, and the testing apparatus.

This thesis has provided evidence that linear relationships work well to describe human colour differentiation abilities. This thesis has also provided evidence that constructing a model of human colour differentiation using these linear relationships can result in accurate predictions about colour differentiation.

9.3 Future Work

The initial experience of developing this modeling system, as well as its successful prediction of differentiation limits for random colours has led to a number of possibilities and issues for future work. Some of these are presented now.

9.3.1 Improvements to the Linear Model

In Chapter 7, a study was performed in which repeated measures of differentiation limits for identical situations were taken. Only two repetitions were performed in this study. It would be very instructive to perform many more repetitions to establish a clear picture of the variations that occur when an individual tries to judge the point of difference. This would provide a deeper understanding of intra-participant variation and influence the accuracy of the predictor.

In Chapter 6, a study was performed to evaluate the effectiveness of using linear functions to model colour differentiation. Future work in this area would include gathering many more data points to further assess the use of linear functions for describing colour differentiation. Gathering many more measurements would also allow the exploration of different functions to define this relationship, such as quadratic or piece-wise linear functions.

Chapter 5 described the calibration process for the model of colour differentiation. To establish each linear function, only endpoints of the line were measured. To increase the confidence and accuracy of the predictor, more points could be collected along the line (between the endpoints). All of the points gathered could be used to generate a line of best fit for them. This should increase the accuracy of the resulting predictor.

9.3.2 Background Colour

An examination into the background colour would help indicate how much the background colour influences colour differentiability. The choice of background colour influences the design of an information visualization in terms of aesthetics and the other colours to be used. As a result, the background colour cannot be assumed to be black as it was for this thesis. Chapter 7 could be performed again using a variety of different colours for the background. To establish whether the background effects colour differentiability, the results presented in this thesis using a black background could be compared to the results of another study using a white background.

9.3.3 Field of Circles

Another interesting set of studies can be formed around the field of circles used in all participant interactions in this thesis. Colour of the circles was the area of interest for this research, but several other properties of the circles could also be examined to determine any effect they have on colour differentiation.

The spacing between circles in this interface was chosen to allow rapid switching focus between two colours. This is why the circles are placed somewhat close together. It would be interesting to see how small and large separations between the circles influence colour differentiation. As the circles get closer, it would seem that the brightness difference between the circles could be relied upon increasingly to differentiate. If this is the case, then it would be reasonable to assume that red and green channel differences would reduce as the circles were brought together, as red and green contribute largely to the perception of brightness. Blue channel differences would probably not decrease though, as this channel contributes little to the perception of brightness. For large distances, memory of colour would begin to influence the results.

9.3.4 Context Switching

When the colour perception context changes, the model presented in this thesis may no longer produce accurate predictions. This is because the model incorporates the factors that influence colour differentiation. This is achieved through calibration of the model using judgement tasks performed inside this context. When the context changes, a new calibration may be necessary.

A comparison study to determine the influence of context changes on the model would be interesting. This would involve studying the affect of varying environmental factors such as ambient lighting, monitor quality, graphics hardware, as well as factors such as fatigue, filters such as glasses or contact lenses, and the time of day. Gathering this information would provide some guidelines as to how frequently a calibration must be performed. It would also provide some ideas regarding the feasibility of performing a contextually-comprehensive set of calibrations once, and then using this set of calibrations to generate a model that most adequately suits the current context. This would greatly reduce the effort needed to use the system.

9.3.5 Additional Colour Tasks

This thesis explored modeling colour differentiation in the restricted context of categorical encoding. It would be illustrative to explore how this system extends to additional colour use techniques in information visualization. These extensions would include colour techniques such as popout, highlighting, brushing, continuums, false colour, and multi-dimensional data visualizations.

9.3.6 Comparison with Traditional Model

The traditional model of colour perception currently used in colour adaptation applications needs to be compared with the model presented in this thesis. It would be informative to see if and when the traditional model performance is superior, and when the model presented in this thesis performs better. This work would serve as a guide for the hybrid approach of combining both of these models to provide maximum benefit to adaptation systems. This could be accomplished by identifying in what circumstances the traditional model fails, and then exploring how the model presented in this thesis can be extended to fulfill this need.

9.3.7 Integration into Existing CATs

Finally, I did not test the model presented in this thesis with actual information visualization techniques. it would be interesting to see how well the system proposed in this thesis can be integrated into existing adaptation software and used on actual information visualizations that use colour to encode information. As these tools are constructed to use the traditional model, redesign of the adaptation systems would be necessary. This would be dependent on the current approach used to interface with the traditional model.

Once this integration was complete, some evaluation of the overall effectiveness of the model presented in this thesis would be useful. This evaluation would examine how well this model serves to assist individuals in atypical colour perception situations using colour-encoded information visualizations. This could be extended to evaluating the performance of the hybrid system described above.

References

[1] N. Alm, A. Astell, G. Gowans, R. Dye, M. Ellis, P. Vaughan, and A. F. Newell. An interactive entertainment system usable by elderly people with dementia. In *Proceedings - HCI 2007*, pages 617–625, Beijing, China, 2007.

[2] Geoffrey Arden, Kemal Gündüz, and Scott Perry. Color vision testing with a computer graphics system: Preliminary results. *Documenta Ophthalmologica*, 69(2):167–174, June 1988.

[3] Leon Barnard, Ji Soo Yi, Julie A. Jacko, and Andrew Sears. Capturing the effects of context on human performance in mobile computing systems. *Personal Ubiquitous Computing*, 11(2):81–96, 2007.

[4] J. Birch, J. L. Barbur, and A. J. Harlow. New method based on random luminance masking for measuring isochromatic zones using high resolution colour displays. *Ophthalmic Physiol Opt.*, 12(2):133–136, April 1992.

[5] Jennifer Birch. *Diagnosis of Defective Colour Vision*. Butterworth Heinemann, Linacre House, Jordan Hill, Oxford, second edition, 2001.

[6] Jennifer Birch. Extreme anomalous trichromatism. In J.D. Mollon, J. Pokorny, and K. Knoblauch, editors, *Normal and Defective Colour Vision*, chapter 38, pages 364–369. Oxford University Press, New York, NY, USA, 2003.

[7] Amy Bruckman and Alisa Bandlow. *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*, chapter Human-Computer Interaction for Kids, pages 428–440. Lawrence Erlbaum Associates, Inc., 2003.

[8] A. Carmichael, A. F. Newell, and M. Morgan. The efficacy of narrative video for raising awareness in ICT designers about older users' requirements. *Interacting with Computers*, 19:587–596, 2007.

[9] Justine Cassell. *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*, chapter Genderizing Human-Computer Interaction, pages 401–412. Lawrence Erlbaum Associates, Inc., 2003.

[10] Barry L. Cole. The handicap of abnormal colour vision. *Clinical and Experimental Optometry*, 87(4-5):258–275, July 2004.

[11] Sara J. Czaja and Chin Chin Lee. *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*, chapter Designing Computer Systems for Older Adults, pages 413–427. Lawrence Erlbaum Associates, Inc., 2003.

[12] Jeremy Daalder. Monitor calibration devices. Website, December 2008. http://www.imagescience.com.au/ColourControl/colourProducts/.

[13] Stephen J. Dain. Clinical colour vision tests. *Clinical and Experimental Optometry*, 87(4-5):276–293, July 2004.

[14] Oculus Opthalmic Diagnostic Devices. Heidelberg multi-color anomaloscope. Website, December 2008. http://www.oculus.de/us/sites/detail_ger.php.

[15] A. Dickinson, M. Smith, J. L. Arnott, A. F. Newell, and R. Hill. Approaches to web search and navigation for older computer novices. In *Proceedings - CHI 2007*, San Jose, CA, USA, 2007. ACM.

[16] Alistair D. N. Edwards, editor. *Extra-Ordinary Human-Computer Interaction*. Cambridge Series on Human-Computer Interaction. Cambridge University Press, New York, New York, USA, 1995.

[17] Dean Farnsworth. Farnsworth-munsell 100-hue and dichotomous tests for color vision. *Journal of the Optical Society of America (1917-1983)*, 33, October 1943.

[18] Brand Fortner and Theodore E. Meyer. *Number by colors: a guide to using color to understand technical data*. Springer-Verlag New York, Inc., New York, NY, USA, 1996.

[19] Julie Fraser and Carl Gutwin. A framework of assistive pointers for low vision users. In *Proceedings - ASSETS 2000*. ACM, 2000.

[20] Tom Fraser and Adam Banks. *Designer's Color Manual: The Complete Guide to Color Theory and Application*. Chronicle Books, San Francisco, California, USA, 2004.

[21] Christopher G. Healey. Choosing effective colours for data visualization. In *VIS '96: Proceedings of the 7th conference on Visualization '96*, pages 263–270, Los Alamitos, CA, USA, 1996. IEEE Computer Society Press.

[22] P. F. Heard, C. J. Stone, R. L. Gregory, and V. J. Marmion. A new computer graphics test for red/green colour anomaly. In G. Verriest, editor, *Colour Vision Deficiencies VIII*, pages 181–194. Martinus Nijhoff, 1987.

[23] Donald Hearn and M. Pauline Baker. *Computer Graphics with OpenGL*. Pearson Prentice Hall, Upper Saddle River, NJ, USA, third edition, 2004.

[24] Ewald Hering. *Outlines of a Theory of the Light Sense*. Harvard University Press, Cambridge, Massachusetts, USA, 1964. Translated by L. Hurvich and D. Jameson.

[25] Hewett, Baecker, Card, Carey, Gasen, Mantei, Perlman, Strong, and Verplank. ACM SIGCHI curricula for human-computer interaction. Website, December 2008. http://sigchi.org/cdg/cdg2.html.

[26] E. Hoadley. Investigating the effects of color. *Communications of the ACM*, 33(2):120–125, 1990.

[27] Jia-Bin Huang, Yu-Cheng Tseng, Se-In Wu, and Sheng-Jyh Wang. Information preserving color transformation for protanopia and deuteranopia. *IEEE Signal Processing Letters*, 14(10):711–714, October 2007.

[28] Gennaro Iaccarino, Delfina Malandrino, Marco Del Percio, and Vittorio Scarano. Efficient edge-services for colorblind users. In *WWW '06: Proceedings of the 15th international conference on World Wide Web*, pages 919–920, New York, NY, USA, 2006. ACM Press.

[29] Gennaro Iaccarino, Delfina Malandrino, and Vittorio Scarano. Personalizable edge services for web accessibility. In *W4A: Proceedings of the 2006 international cross-disciplinary workshop on Web accessibility (W4A)*, pages 23–32, New York, NY, USA, 2006. ACM Press.

[30] M. Ichikawa, K. Tanaka, S. Kondo, K. Hiroshima, K. Ichikawa, S. Tanabe, and K. Fukami. Web-page color modification for barrier-free color vision with genetic algorithm. In *Proceedings - Genetic and Evolutionary Computation - GECCO 2003*, volume 2724 of *Lecture Notes in Computer Science*, pages 2134–2146, Berlin/Heidelberg, Germany, 2003. Springer.

[31] M. Ichikawa, K. Tanaka, S. Kondo, K. Hiroshima, K. Ichikawa, S. Tanabe, and K. Fukami. Preliminary study on color modification for still images to realize barrier-free color vision. *2004 IEEE International Conference on Systems, Man and Cybernetics*, 1:36–41, October 2004.

[32] Shinobu Ishihara. Tests for colour-blindness. Pseudoisochromatic Plates, 1917.

[33] Julie A. Jacko, Holly S. Vitense, and Ingrid U. Scott. *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*, chapter Perceptual Impairments and Computing Technologies, pages 504–522. Lawrence Erlbaum Associates, Inc., 2003.

[34] Luke Jefferson and Richard Harvey. Accommodating color blind computer users. In *ASSETS '06: Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility*, pages 40–47, New York, NY, USA, 2006. ACM.

[35] Luke Jefferson and Richard Harvey. An interface to support color blind computer users. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 1535–1538, New York, NY, USA, 2007. ACM.

[36] Tanja Keller, Peter Gerjets, Katharina Scheiter, and Barbel Garsoffky. Information visualizations for knowledge acquisition: The impact of dimensionality and color coding. *Computers in Human Behavior: Instructional Design for Effective and Enjoyable Computer-Supported Learning*, 22(1):43–65, January 2006.

[37] Rae Kokotailo and Donald Kline. Colour perception in everyday life. Website, December 2008. http://www.psych.ucalgary.ca/PACE/VA-LAB/colourperceptionweb/congenital.htm.

[38] Alexander D. Logvinenko and Lesley L. Beattie. Partial colour matching: a new method to evaluate colour appearance and derive unique hues. In *Proceedings - CGIV 2006 - Third European Conference on Color in Graphics, Imaging and Vision*, Leeds, UK, 2006. Society for Imaging Science and Technology.

[39] Aaron Marcus. *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*, chapter Global and Intercultural User-Interface Design, pages 441–463. Lawrence Erlbaum Associates, Inc., 2003.

[40] John D. McFall. Managing and reproducing colour images (MARCI). In *CASCON '93: Proceedings of the 1993 conference of the Centre for Advanced Studies on Collaborative research*, pages 1164–1179, Toronto, Ontario, Canada, 1993. IBM Press.

[41] ICC Members. The role of ICC profiles in a colour reproduction system. White Paper 7, International Color Consortium, December 2004.

[42] J. Nam, Yong Man Ro, Y. Huh, and M. Kim. Visual content adaptation according to user perception characteristics. *IEEE Transactions on Multimedia*, 7(3):435–445, June 2005.

[43] M. Neitz and J. Neitz. Numbers and ratios of visual pigment genes for normal red-green color vision. *Science*, 267(5200):1013–1016, February 1995.

[44] A. Newell, J. Arnott, A. Carmichael, and M. Morgan. Methodologies for involving older adults in the design process. In *Proceedings - HCI 2007*, pages 982 – 989, Beijing, China, 2007.

[45] Alan F. Newell. *Extra-Ordinary Human Computer Interaction*, chapter Introduction, pages 3–18. Cambridge University Press, 1995.

[46] Alan F. Newell, Alex Carmichael, Peter Gregor, and Norman Alm. *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*, chapter Information Technology for Cognitive Support, pages 464–481. Lawrence Erlbaum Associates, Inc., 2003.

[47] Noboru Ohta and Alan Robertson. *Colorimetry: Fundamentals And Applications*. John Wiley & Sons, Wiley-IS&T Series in Imaging Science and Technology, 2005.

[48] Karl Rasche, Robert Geist, and James Westall. Detail preserving reproduction of color images for monochromats and dichromats. *IEEE Computer Graphics and Applications*, 25(3):22–30, 2005.

[49] Karl Rasche, Robert Geist, and James Westall. Re-coloring images for gamuts of lower dimension. *Computer Graphics Forum*, 24(3):423–432, September 2005.

[50] B. C. Regan, J. P. Reffin, and J. D. Mollon. Luminance noise and the rapid determination of discrimination ellipses in colour deficiency. *Vision Research*, 34(10):1279–1299, May 1994.

[51] Yong Man Ro and Seungji Yang. Color adaptation for anomalous trichromats. *International Journal of Imaging Systems and Technology*, 14(1):16–20, 2004.

[52] P.K. Robertson and J.F. O'Callaghan. The generation of color sequences for univariate and bivariate mapping. *IEEE Computer Graphics and Applications*, 6(2):24–32, February 1986.

[53] Oliver Sacks. *The Island of the Colorblind*. Vintage, United Kingdom, January 1998.

[54] Harvey Richard Schiffman. *Sensation and Perception: An Integrated Approach*. John Wiley & Sons, Inc., New York, NY, USA, fifth edition, 2001.

[55] Andrew Sears and Mark Young. *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*, chapter Physical Disabilities and Computing Technologies: An Analysis of Impairments, pages 482–503. Lawrence Erlbaum Associates, Inc., 2003.

[56] Ben Shneiderman. *Designing the User Interface*. Addison-Wesley, Reading, Massechusetts, USA, third edition, 1998.

[57] Alvy Ray Smith. Color gamut transform pairs. In *SIGGRAPH 78 Conference Proceedings*, pages 12–19, Atlanta, Georgia, USA, 1978. Association for Computing Machinery.

[58] Jaeil Song, Seungji Yang, Cheonseog Kim, Jaeho Nam, Jin-Woo Hong, and Yong Man Ro. Digital item adaptation for color vision variations. In Bernice E. Rogowitz and Thrasyvoulos N. Pappas, editors, *Human Vision and Electronic Imaging VIII*, volume 5007, pages 96–103. SPIE, 2003.

[59] Maureen C. Stone. *A Field Guide to Digital Color*. A. K. Peters, Natick, Massachusetts, USA, 2003.

[60] Maureen C. Stone. Color in information display: Principles, perception, and models, 2004. Published as part of Course 20 at SIGGRAPH 2004.

[61] Luigi Troiano, Cosimo Birtolo, and Maria Miranda. Adapting palettes to color vision deficiencies by genetic algorithm. In *GECCO '08: Proceedings of the 10th annual conference on Genetic and evolutionary computation*, pages 1065–1072, New York, NY, USA, 2008. ACM.

[62] Edward R. Tufte. *Envisioning Information*. Graphics Press, Cheshire, Connecticut, USA, tenth edition, 1990.

[63] Gregg Vanderheiden. *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*, chapter Interaction For Diverse Users, pages 397–400. Lawrence Erlbaum Associates, Inc., 2003.

[64] F. Viénot, H. Brettel, L. Ott, A. Ben M'Barek, and J. D. Mollon. What do colour-blind people see? *Nature*, 376:127–128, July 13 1995.

[65] Howard Wainer and Carl M. Francolini. An empirical inquiry concerning human understanding of two-variable color maps. *The American Statistician*, 34(2):81–93, May 1980.

[66] Ken Wakita and Kenta Shimamura. Smartcolor: Disambiguation framework for the colorblind. In *Assets '05: Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility*, pages 158–165, New York, NY, USA, 2005. ACM Press.

[67] Brian A. Wandell. *Foundations of Vision*. Sinauer Associates, Inc., Sunderland, Massachusetts, USA, 1995.

[68] Colin Ware. *Information Visualization - Perception for Design*. Morgan Kaufmann Publishers, San Francisco, California, USA, 2000.

[69] Colin Ware and John C. Beatty. Using color dimensions to display data dimensions. *Human Factors*, 30(2):127–142, 1988.

[70] Seungji Yang and Yong Man Ro. Visual contents adaptation for color vision deficiency. In *Proceedings - International Conference on Image Processing - 2003*, pages 453–456. IEEE, 2003.

[71] Seungji Yang, Yong Man Ro, Jeho Nam, Jinwoo Hong, Sang Yul Choi, and Jin-Hak Lee. Improving visual accessibility for color vision deficiency based on MPEG-21. *ETRI Journal*, 26(3):195–202, June 2004.

Appendix A

Pre-Study Questionnaire

Demographics survey - Participant #________ - Model Study #1

1. Age: ________

2. Sex: **M** **F** (circle one)

3. Occupation: ____

4. If you are a student, list your level, university, major, and year:

____

5. How many hours a week, on average, do you spend working with computers?

____

6. List 5 applications that you use frequently (e.g., Firefox, Internet Explorer, Outlook):

i. ____________ ii. ____________ iii. ____________ iv. ___________ v. ___________

7. How many hours a week, on average, do you spend playing electronic games (e.g., Wii, XBox, PlayStation, computer games, World of Warcraft, Enemy Territory)?

____

8. Please list a few of the game(s) you play most frequently:

____

9. Do you have any experience with information visualizations? **Y** **N** (circle one)

10. If yes, please list some visualizations and visualization tools you have used:

____

11. Please rank your (corrected) visual acuity:

excellent good fair poor

12. Have you been previously diagnosed with colour blindness? **Y** **N** (circle one)

13. To the best of your knowledge, do any of the following relatives have colour blindness?

mother's father: **Y** **N** (circle one)

mother's brother(s): **Y** **N** (circle one)